CONSTRAINED MEASUREMENT SYSTEMS OF LOW-DIMENSIONAL SIGNALS

A Thesis
Presented to
The Academic Faculty

by

Han Lun Yap

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
May 2013

CONSTRAINED MEASUREMENT SYSTEMS OF LOW-DIMENSIONAL SIGNALS

Approved by:

Professor Christopher J. Rozell,
Advisor
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Professor Mark A. Davenport
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Professor Robert J. Butera
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Professor Mohammad Ghomi
School of Mathematics
Georgia Institute of Technology

Professor Justin K. Romberg
School of Electrical and Computer
Engineering
Georgia Institute of Technology

Professor Michael B. Wakin
Department of Electrical Engineering
and Computer Science
Colorado School of Mines

Date Approved: 9 November 2012

ACKNOWLEDGEMENTS

Thanks to generous funding by DSO National Laboratories of Singapore, I am very fortunate to be able to pursue my PhD education at Georgia Tech. Thus, Georgia Tech has become a home (very) far away from home for the past four and a half years, and my wife and I are grateful to many friends and colleagues that have made our stay the most wonderful possible.

I will first like to acknowledge my committee members for their time and energy throughout both my proposal and thesis processes. In particular, I will always be grateful to my advisor Chris Rozell for his commitment to my academic and professional maturity. I have learned alot from him, benefiting not only from his immense knowledge, but also from his insistence on perfecting one's presentation skills. I am also grateful to Justin Romberg for being my "surrogate" advisor, allowing me to sit in his group meetings and later joining them for beers. Mike Wakin has always been involved in my PhD education as a collaborator in many of my projects. From him, I have learned to be meticulous in my work (especially in lengthy math proofs!) and have definitely benefitted from his vast knowledge. As a student in his Riemannian geometry classes, Mohammad Ghomi expanded my Euclidean world to include manifolds. Besides his insightful lectures, I've also benefitted tremedously from our discussions outside of class. Sitting just next door to my advisor (in his previous office), Robert Butera has in many occasions been exceptionally helpful in providing valuable insights into problems related to neurons and dynamical systems. Last but not least, I will like to thank Mark Davenport for all the stimulating discussions we had when we met at various conferences before he came to Georgia Tech.

Graduate school would be dull if not for all my "colleagues". Thus, I will like

to thank Adam (and Diane), Mengchen, Abbie, Aurèle, Aditya, Darryl (and Anne), Chris Turnes, Ali, Salman, William, Dhruv, Allie, Nick, and Andrew for their companionship and stimulating discussions (and beers!) inside and outside of Chris's and Justin's group meetings. I will miss the days of going to the "Faculty Dining Room" after our group meetings. Not to forget my collaborators Jae and Armin, with whom I have spent many (late) hours dotting the i's and crossing the t's in our papers. Many thanks also go out to my friends outside of work: Sean, Leo, Jeong, Seksan, Borhan, Alexandro, Andrew Lim, Zhou Min and Ling Zhi, John Poh, and Darren and Danessa Foo (and little Jasmin), who have made my stay here in the USA more comfortable and infinitely more entertaining.

Finally, I will not be where I am today if not for the people closest to me. First, I will like to thank my mom for guiding me to where I am today. The most important of all, I will like to thank my wife, Ailian, for always providing all her love, support, encouragements, and distractions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The object of this thesis is the study of constrained measurement systems of signals having low-dimensional structure using analytic tools from Compressed Sensing (CS). Realistic measurement systems usually have architectural constraints that make them differ from their idealized, well-studied counterparts. Nonetheless, these measurement systems can exploit structure in the signals that they measure. Signals considered in this research have low-dimensional structure and can be broken down into two types: static or dynamic. Static signals are either sparse in a specified basis or lying on a low-dimensional manifold (called manifold-modeled signals). Dynamic signals, exemplified as states of a dynamical system, either lie on a low-dimensional manifold or have converged onto a low-dimensional attractor.

In CS, the Restricted Isometry Property (RIP) of a measurement system ensures that distances between all signals of a certain sparsity are preserved. This stable embedding ensures that sparse signals can be distinguished one from another by their measurements and therefore be robustly recovered. Moreover, signal-processing and data-inference algorithms can be performed directly on the measurements instead of requiring a prior signal recovery step. Taking inspiration from the RIP, this research analyzes conditions on realistic, constrained measurement systems (of the signals described above) such that they are stable embeddings of the signals that they measure.

Specifically, this thesis focuses on four different types of measurement systems. First, we study the concentration of measure and the RIP of random block diagonal matrices that represent measurement systems constrained to make local measurements. Second, we study the stable embedding of manifold-modeled signals by

existing CS matrices. The third part of this thesis deals with measurement systems of dynamical systems that produce time series observations. While Takens' embedding result ensures that this time series output can be an embedding of the dynamical systems' states, our research establishes that a stronger stable embedding result is possible under certain conditions. The final part of this thesis is the application of CS ideas to the study of the short-term memory of neural networks. In particular, we show that the nodes of a recurrent neural network can be a stable embedding of sparse input sequences.

CHAPTER I

INTRODUCTION

A primary task of man-made and biological sensory systems is to provide an efficient and stable representation of the signals they are measuring. These representations are then acted upon by higher-level processing units to extract useful information. Unsurprisingly, the stability of the representation will determine the quality of the information extracted and reduce the susceptibility of this extraction process to noise. On the other hand, the efficiency of the representation will dictate the energy dispensed by the measurement process and the eventual storage requirement for the information measured.

Over the course of human history, man-made sensors have made great strides in reproducing and sometimes surpassing the capabilities of their biological counterparts. For example, airborne radar systems can penetrate clouds to provide high-resolution images of a ground scene (synthetic aperture radars), telescopes are sensitive enough to detect exoplanets (NASA's Kepler mission), and cameras can capture enough information to allow focusing and refocusing on any part of an image (light-field cameras). All of these examples are made possible by the incessant march of technological innovation, particularly that coming from improvements in resolution, higher bit-rate recordings, and new sensor modalities. With the economy of scale driving sensor prices down, man-made sensors have also proliferated modern society. The resulting rapid increase in sensor data has now exceeded our storage and processing capabilities [12]. As a result of this tremendous amount of data collected by sensors, it has become increasingly important to think, not only about new ways of sensing, but also about integrating processing with sensing to conserve resources.

1.1 The Compressed Sensing Revolution

The emerging field of compressed sensing (CS) has brought about a change in the way we think about the seemingly separate notions of sensing and compression [25]. Compressing high-dimensional data down to a more manageable size with a minimal loss of information is one way to make maximum use of limited storage and communication resources. However, traditional compression schemes are suboptimal in that we need to acquire the high-dimensional signal before throwing away most of the bits via compression. CS brought about a paradigm shift by showing that we can do both the compression and the sensing step at the same time. The main results of CS show that measurement systems can be designed to provide efficient and stable representations of signals with data sizes that scale only with the information content of the signal.

To achieve this compression, CS relies on a signal model to make up for the lack of measurements. The signal model used in CS is a sparsity model, meaning that signals are assumed to be made up of linear combinations of a few atoms in a dictionary. Many natural signals have approximately sparse representations in some dictionaries, and this fact has been exploited in many compression algorithms (e.g., JPEG and MP3 [94]). The breakthrough result of CS shows that whenever the sparsity structure of the signal is maintained in the measurement space (i.e., the measurements form a stable representation of the signal), a convex and tractable optimization program can be used to recover the measured signal even under noisy conditions.

Seen from another perspective, the preservation of the signal's sparsity structure implies that full information about the signal is maintained in the measurement space. In CS, this sufficient condition for information preservation is called the restricted isometry property (RIP). The RIP (discussed in detail in Section 2.4) is a precise technical characterization of when a measurement process preserves the geometric structure of the signal family being sensed, and thus provides a stable representation

of the information. The surprising fact is that the number of measurements needed for a measurement system to satisfy the RIP can scale linearly with the sparsity of the signal and only logarithmically with its dimension. Not only does the RIP guarantee the robust recovery of signals from underdetermined measurements, but it also guarantees the performance of various data-processing and inference algorithms in the compressed space without requiring full signal reconstruction [43].

1.2 Signal Models and their Representations

Despite the success of sparsity models, some signal families are more naturally characterized by other types of low-dimensional geometric structures. For one example, some signal families are more appropriately described as lying on a low-dimensional manifold. An example of such a family is a collection of different poses of an object (see Figure 1). Although it may take a lot of pixels to faithfully represent a single pose of the object, the whole family of images can be parameterized by the few degrees of freedom that distinguish the different poses. When viewed as a geometrical object in the high-dimensional image space, the whole family of images lie on (or close to) a low-dimensional manifold. Surprisingly, the RIP idea in CS has been extended to measurement systems for signal families lying on manifolds. Indeed, a measurement system represented by a random orthoprojector matrix can be a stable embedding of a manifold as long as the number of measurements is proportional to the dimension of the manifold [15]. Similar to the RIP, a stable embedding implies that Euclidean (or geodesic) distances are approximately maintained between points in the measurement domain. Again, even though the number of measurements can be much smaller than the ambient dimension of the signal, key properties of the manifold are retained in the measurement domain. Exploiting this fact, manifold learning algorithms such as ISOMAP [129] can run more efficiently in the low-dimensional

measurement space while suffering little loss in accuracy. Just as in CS, manifold-modeled signals can be robustly recovered from their measurements using an iterative projection algorithm [122] whenever the measurement system is a stable embedding of the manifold.

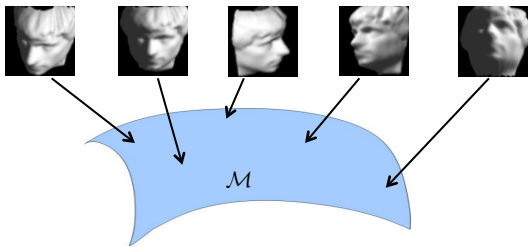


Figure 1: Images of different poses of a head seen as lying on a low-dimensional manifold in a high-dimensional space.

For another example, dynamic signals can have low-dimensional structure in their temporal evolution. Dynamic signals can be modeled by the states of a dynamical system, meaning that a deterministic equation (e.g., an ordinary differential equation) governs the evolution of the signal over time. When the dynamical system is dissipative, the system converges onto a low-dimensional (chaotic) attractor manifold. Measurement systems of such signals have a temporal constraint; they only have a short window of time to take measurements of the signal before it evolves. For example, a video sequence is a dynamic signal where only a single frame of the video is seen by a measurement system at any time. Given this temporal constraint, an interesting question to ask is whether measurement systems of dynamic signals can still form a meaningful representation of the signals. A surprising result by Floris Takens [126] showed that a time-series observation of the system states can be an embedding of the attractor that governs the behavior of the dynamic system. An embedding ensures that the topology of the signal family is preserved in the measurement space, meaning that no two states result in the same time-series output. Although an embedding is weaker than a stable embedding because distances between points are not necessarily preserved, this result has been exploited in many time-series algorithms (e.g.,

time-series prediction, noise reduction, and nonlinear systems identification [81]).

1.3 Structured Sensing Systems

The prototypical measurement system known to satisfy the RIP (with high probability) is a random matrix whose entries are independent and identically distributed Gaussian random variables. Although this Gaussian random matrix has many nice analytic properties, it has three glaring drawbacks that prevent its widespread use in practice. First, no efficient way exists to compute the multiplication of this matrix with a vector, resulting in algorithms with high computational complexity. Second, this type of large random matrix is difficult to store in the memory of most measurement systems. Third, a Gaussian random matrix cannot account for many of the architectural constraints seen in practical sensing systems.

As a result of these practical issues, there has been increasing interest in demonstrating the RIP (or generally a stable embedding) for structured random matrices that overcome one or more of the above-mentioned issues. Examples of structured matrices studied in the literature include subsampled Fourier matrices [118] and subsampled circulant or Toeplitz matrices [83, 112]. Both of these matrix types have fast transforms, efficient storage in memory, and are models of physical sensing systems (e.g., medical imaging systems, sensor arrays, and radars). However, many other sensing systems have constraints that make it unclear whether their measurements can be a stable representation for the signal class of interest.

1.4 Contributions

The focus of this thesis is on deriving conditions whereby realistic, structured measurement systems can provide a stable embedding of various signal families. Once stable embedding is established, we can then reap benefits from the various recoverability and signal-processing guarantees in the existing literature.

The first part of our work (presented in Chapter 3) deals with measurement systems having an architectural constraint that prevents global data aggregation. These constraints can be spatial or temporal (e.g., distributed sensor arrays and dynamically changing signals), and such measurement systems can be modeled by block diagonal matrices. In this work, we show that block diagonal matrices have concentration of measure inequalities that depend on the type of signals they are measuring. Furthermore, we also show that these matrices can satisfy the RIP with measurement rates that depend on the sparsity basis appropriate for the signals.

The second focus of our work (presented in Chapter 4) deals with measurement systems of the more general class of manifold-modeled signals. In this work, we expand the class of measurement systems that can provide stable embeddings of manifolds to include systems that possess fast transforms or having various architectural constraints (e.g., systems represented by block diagonal matrices discussed above).

Our third contribution (presented in Chapter 5) is in the area of stable embeddings for time-series analysis. As discussed previously, Takens' embedding theorem only ensures an embedding of the system attractor in the time-series measurements, and an embedding is prone to corruption by measurement noise. Here, we expand on Takens' theorem by showing conditions whereby the time-series output of a dynamical system can be a stable embedding of the system attractor.

Finally (in Chapter 6), we turn our focus to more biologically inspired measurement architectures. In this work, we show that a recurrent neural network can retain a working memory of its input in the transient network state. In particular, we show that the network states can be a stable embedding of sparse input signals. Furthermore, this stable embedding can be achieved with a number of neurons that is very much less than the length of the input signal.

The results in this thesis are not mine alone but are the outcomes of very fruitful collaborations with many other researchers. Their respective contributions will be

highlighted in the opening paragraphs of each of the relevant chapters.

CHAPTER II

BACKGROUND

In this chapter, we survey the literature for relevant work on stable representations of various signal classes and the guarantees that these representations provide for various algorithms. We start by presenting in Section 2.1 the mathematical notations and definitions used throughout this thesis. In Section 2.2, we introduce the notions of embeddings and stable embeddings that will be used to quantify the stability of a representation. Then, Section 2.3 focuses on the stable embeddings of finite signal families known as the Johnson-Lindenstrauss lemma in the computer science community. Following this, Section 2.4 gives an overview of Compressed Sensing (CS) and introduces the concept of the Restricted Isometry Property (RIP), where the RIP of a measurement system ensures the stable embedding a sparse signal family. Finally, Section 2.5 discusses the embeddings and stable embeddings of manifold-modeled signal families.

2.1 Mathematical Preliminaries

2.1.1 Vectors and Matrices

For $x = [x_1, \cdots, x_N]^T \in \mathbb{C}^N$, denote the complex[1] variable by j, the (element-wise) complex conjugate by x^*, and the Hermitian transpose by $x^H = (x^*)^T$. $\mathrm{Re}\{x\} \in \mathbb{R}^N$ is the (element-wise) real part of x and $\mathrm{Im}\{x\} \in \mathbb{R}^N$ is the (element-wise) imaginary part of x. For $p \geq 1$, denote by $\|x\|_p := \left(\sum_{i=1}^N |x_i|^p\right)^{1/p}$ the ℓ_p-norm of x. We define $\|x\|_\infty := \max_n |x_n|$ as the limiting ℓ_p-norm of x as $p \to \infty$ and $\|x\|_0$ as the counting

[1]For generality, we use complex numbers in this section. Most signals that we encounter in this thesis are real-valued signals, and the discussions here will hold equally well over the real field.

pseudo-norm[2] that returns the number of non-zero entries of x.

For a matrix $\Phi \in \mathbb{C}^{M\times N}$, denote by $\|\Phi\|_2$ the spectral norm of Φ (i.e., the largest absolute singular value of Φ) and by $\|\Phi\|_F$ its Frobenius norm (i.e., the ℓ_2-norm of the singular values of Φ or equivalently the square-root of the sum of absolute-value squared of all entries in Φ).

Define the unit-norm operator $U : \mathbb{C}^N - \{0\} \to \mathbb{S}^{N-1}$ that takes a non-zero vector and projects it onto the unit sphere (i.e., for any $x \in \mathbb{C}^N - \{0\}$, $U(x) := \frac{x}{\|x\|_2}$). Similarly, U can also act on a non-zero matrix Φ by normalizing it by its Frobenius norm (i.e., $U(\Phi) := \frac{\Phi}{\|\Phi\|_F}$).

For a vector $x \in \mathbb{C}^N$, we define the diagonal operator, $\text{diag} : \mathbb{C}^N \to \mathbb{C}^{N\times N}$, as the construction of a $N \times N$ diagonal matrix $\text{diag}(x)$ with its diagonal entries corresponding to the vector x.

2.1.2 Sets and Manifolds

For any integer J, denote $[J] := \{1, 2, \cdots, J\}$ (i.e., $[J]$ is the list of integers from 1 to J). For any subset $\mathcal{S} \subset \mathbb{C}^N$, we define the action of the unit-norm operator U on $\mathcal{S}$ as $U(\mathcal{S}) := \{U(x) \mid x \in \mathcal{S} - \{0\}\}$. Next, we define the difference between any two subsets $\mathcal{S}_1, \mathcal{S}_2 \subset \mathbb{C}^N$ as the set comprised of pairwise differences between the elements of the sets, $\mathcal{S}_1 - \mathcal{S}_2 := \{a - b \mid a \in \mathcal{S}_1,\ b \in \mathcal{S}_2\}$. Combining the two definitions above, we see that $U(\mathcal{S} - \mathcal{S}) = \{U(a - b) \mid a, b \in \mathcal{S},\ a \neq b\}$ for some set $\mathcal{S}$. For a finite subset $\mathcal{S} \subset \mathbb{C}^N$, $|\mathcal{S}|$ denotes the cardinality of the set. If $\mathcal{S}$ is an infinite set, let $\text{vol}(\mathcal{S})$ denote its volume defined as $\text{vol}(\mathcal{S}) = \int_{\mathcal{S}} dv$, where dv is the volume element on $\mathcal{S}$.

A *manifold* of dimension K is a (second countable, Hausdorff) topological space that is locally homeomorphic to the Euclidean space $\mathbb{C}^K$ [52]. To say a manifold is locally homeomorphic to $\mathbb{C}^K$ means that for every point on the manifold, we can find a neighborhood U around it that can be mapped by a chart $\phi : U \to \mathbb{C}^K$ in

[2]For $p < 1$, $\|\cdot\|_p$ no longer satisfies the triangle inequality and therefore becomes a pseudo-norm (instead of a norm).

an invertible manner to an open K-dimensional Euclidean ball. Manifolds that are embedded in a (larger) Euclidean space are called *submanifolds*. A manifold is called smooth if for any overlapping neighborhoods U_1, U_2 with their corresponding charts ϕ_1, ϕ_2, the transition map defined as $\phi_2 \circ \phi_1^{-1} : \phi_1(U_1 \cap U_2) \to \phi_2(U_1 \cap U_2)$ is a smooth (i.e., infinitely differentiable) function. Then, a *Riemannian* manifold is defined as a smooth manifold that is further endowed with a Riemannian metric. Having a metric allows us to define lengths of paths between points on a Riemannian manifold.

For any manifold $\mathcal{M}$, we let $\mathcal{T}_x\mathcal{M}$ denote the tangent space of $\mathcal{M}$ at any point $x \in \mathcal{M}$. In our analyses, $\mathcal{T}_x\mathcal{M}$ can be thought of as a K-dimensional linear subspace of $\mathbb{C}^N$ passing through the origin. Let $\mathcal{T}\mathcal{M}$ be the *tangent bundle* of the manifold $\mathcal{M}$, defined as the union of all tangent spaces in $\mathcal{M}$ (i.e., $\mathcal{T}\mathcal{M} = \bigcup_{x \in \mathcal{M}} \mathcal{T}_x\mathcal{M}$). A vector field on a manifold $\mathcal{M}$, $X : \mathcal{M} \to \mathcal{T}\mathcal{M}$, is a function that assigns a tangent vector $X(x) \in \mathcal{T}_x\mathcal{M}$ to every point $x \in \mathcal{M}$. We denote by $\mathcal{X}(\mathcal{M})$ the set of all vector fields on $\mathcal{M}$.

If $\mathcal{M}$ is a Riemannian manifold, we define the geodesic distance between two points $x, y \in \mathcal{M}$, $d_{\mathcal{M}}(x, y)$, by the length of the shortest path between x and y along on the manifold. Let $\mathcal{B}^{\mathcal{M}}_{d_{\mathcal{M}}}(x, \epsilon)$ be the geodesic ball centered at $x \in \mathcal{M}$ of radius ϵ (i.e., $\mathcal{B}^{\mathcal{M}}_{d_{\mathcal{M}}}(x, \epsilon) := \{y \in \mathcal{M} \mid d_{\mathcal{M}}(y, x) \leq \epsilon\}$). Similarly, let $\mathcal{B}^K_p(x, \epsilon)$ be the Euclidean ℓ_p-ball of radius ϵ centered at $x \in \mathbb{C}^K$ (i.e., $\mathcal{B}^K_2(x, \epsilon) := \{y \in \mathbb{C}^K \mid \|y - x\|_p \leq \epsilon\}$). We say that a finite set $\mathcal{C}$ is an $(\epsilon, d_{\mathcal{M}})$-*cover* for $\mathcal{M}$ if $\mathcal{M} \subset \bigcup_{b \in \mathcal{C}} \mathcal{B}^{\mathcal{M}}_{d_{\mathcal{M}}}(b, \epsilon)$. This implies that for every $x \in \mathcal{M}$, we can find a $b \in \mathcal{C}$ such that $d_{\mathcal{M}}(b, x) \leq \epsilon$. The $(\epsilon, d_{\mathcal{M}})$-cover $\mathcal{C}$ with the minimal cardinality is denoted by $\mathcal{C}(\mathcal{M}, d_{\mathcal{M}}, \epsilon)$, and the cardinality of $\mathcal{C}(\mathcal{M}, d_{\mathcal{M}}, \epsilon)$ is called the $(\epsilon, d_{\mathcal{M}})$-covering number of $\mathcal{M}$ or simply the geodesic covering number. Similarly, if $\mathcal{M}$ is also a submanifold of $\mathbb{C}^N$, we say that the finite set $\mathcal{C}(\mathcal{M}, \|\cdot\|_2, \epsilon)$ (of minimal cardinality) is an $(\epsilon, \|\cdot\|_2)$-cover $\mathcal{M} \subset \mathbb{C}^N$ if $\mathcal{M} \subset \bigcup_{b \in \mathcal{C}(\mathcal{M}, \|\cdot\|_2, \epsilon)} \mathcal{B}^N_2(b, \epsilon)$.

2.1.3 Random Variables

Subgaussian random variables are often invoked for probabilistic analysis thanks to many convenient properties. We define a subgaussian random variable as follows:

Definition 2.1.1. *[136] A random variable Z is* subgaussian *if there exists a constant $a > 0$ such that*

$$\mathbb{E}\left\{|Z|^p\right\}^{1/p} \leq a\sqrt{p} \text{ for all } p \geq 1.$$

The quantity $\|Z\|_{\psi_2} := \sup_{p\geq 1} p^{-1/2}\mathbb{E}\left\{|Z|^p\right\}^{1/p}$ is called the subgaussian norm of Z.

Examples of subgaussian random variables include zero-mean Gaussian random variables, ± 1 Bernoulli random variables (each value with probability $\frac{1}{2}$), and uniform random variables on the interval $[-1, 1]$. For a given subgaussian random variable Z, the variance $\mathbb{E}\left\{Z^2\right\}$ is a constant multiple of $\|Z\|_{\psi_2}^2$ with $0 \leq \frac{\mathbb{E}\left\{Z^2\right\}}{\|Z\|_{\psi_2}^2} \leq \sqrt{2}$; the exact value of $\frac{\mathbb{E}\left\{Z^2\right\}}{\|Z\|_{\psi_2}^2}$ depends on the specific distribution for Z under consideration (Gaussian, Bernoulli, etc.). In what follows, we use the abbreviation *i.i.d.* to denote the term *independently and identically distributed*. A sequence of i.i.d. Bernoulli random variables is called a *Rademacher* sequence. In this thesis, $\overset{\text{dist.}}{=}$ means that the random variables on both sides of the equality have the same distribution.

2.1.4 Functions and Comparisons

A Lipschitz map $f : \mathbb{C}^N \to \mathbb{C}^M$ is defined as a mapping such that $\|f(x) - f(y)\|_2 \leq C\|x - y\|_2$ for some constant C and all $x, y \in \mathbb{C}^N$. A bi-Lipschitz map $f : \mathbb{C}^N \to \mathbb{C}^M$ is defined as a mapping such that $C_l\|x-y\|_2 \leq \|f(x)-f(y)\|_2 \leq C_u\|x-y\|_2$ for some constant C_l, C_u and all $x, y \in \mathbb{C}^N$. We say that a function f is $\mathcal{C}^k$ for any integer $k > 0$ if its k-th derivative is continuous. Then, we say that a function f is smooth if it is $\mathcal{C}^k$ for all $k > 0$. We denote by $\mathcal{C}^k(\Omega_1, \Omega_2)$ as the space of all $\mathcal{C}^k$ functions $f : \Omega_1 \to \Omega_2$, and we denote the differential operator at any point $x \in \mathcal{M}_1$ as $D_x f$. If $\mathcal{M}_1, \mathcal{M}_2$ are manifolds and $f : \mathcal{M}_1 \to \mathcal{M}_2$ is at least $\mathcal{C}^1$, $D_x f$ can be regarded as

a matrix operating on the vectors in $\mathcal{T}_x\mathcal{M}_1$ to give vectors in $\mathcal{T}_{f(x)}\mathcal{M}_2$. A function $f : \mathcal{M}_1 \to \mathcal{M}_2$ is a homeomorphism (and thus $\mathcal{M}_1$ is homeomorphic to $\mathcal{M}_2$) if f is continuous and has a continuous inverse. A function $f : \mathcal{M}_1 \to \mathcal{M}_2$ is an immersion on $\mathcal{M}_1$ if its derivatives are injective everywhere on $\mathcal{M}_1$ (i.e., $D_x f$ is full-rank for all $x \in \mathcal{M}_1$). Finally, we use the notation $a \gtrsim b$ (or $a \lesssim b$) to mean that there exists a constant C such that $a \geq Cb$ (or $a \leq Cb$ respectively). If the constant C depends on some parameter τ, we write $a \gtrsim_\tau b$ to highlight the dependence on this parameter.

2.2 Embeddings and Stable Embeddings

The focus of this thesis is to show that various sensory systems provide a stable representation of the signals they are measuring. We use the notion of an embedding to quantify the representation of a signal, and the stability of an embedding depends on the conditioning of that embedding. In this section, we make explicit these notions of *embedding* and *conditioning*.

For simplicity, we restrict our attention to finite-dimensional signals in Euclidean space denoted by $x \in \mathbb{R}^N$. A signal family, denoted by $\mathcal{M} \subset \mathbb{R}^N$, is either a finite or infinite collection of signals having some common characteristics (e.g., sparsity or lying on a manifold). Suppose we have a measurement system acting on the signal family $\mathcal{M}$ represented by an operator $\Phi : \mathbb{R}^N \to \mathbb{R}^M$. We say that Φ is a *topological embedding* of $\mathcal{M}$ if Φ is injective on $\mathcal{M}$ and $\mathcal{M}$ is homeomorphic to $\Phi(\mathcal{M})$. If the signal family $\mathcal{M}$ is a Riemannian submanifold of the ambient Euclidean space $\mathbb{R}^N$, then we say that Φ is a *smooth embedding* of $\mathcal{M}$ if Φ is a topological embedding and an immersion on $\mathcal{M}$. We remark that when $\mathcal{M} \subset \mathbb{R}^N$ has a manifold structure, then $\Phi(\mathcal{M})$ is also a submanifold in $\mathbb{R}^M$. We also remark that if Φ is linear, then the two notions of embedding (i.e., topological and smooth) coincide. Essentially, an embedding ensures that distinct points remain distinct in the measurement space and that the structure of the signal family is not destroyed.

However, simply ensuring that a measurement operator $\Phi : \mathbb{R}^N \to \mathbb{R}^M$ provides an embedding of a set $\mathcal{M} \subset \mathbb{R}^N$ may not be sufficient when the measurement process is noisy. Points that were initially well-separated (and easily distinguishable) in the ambient space $\mathbb{R}^N$ may be mapped close together in the measurement space $\mathbb{R}^M$. When the measurements are corrupted by noise, these points may become indistinguishable. Therefore, we may be interested in differentiating measurement systems based on how well they maintain distances between points in a signal family. One way to do so is by considering if measurement systems provides a *stable embedding* of the signal family $\mathcal{M}$.

We say that an operator Φ provides a stable embedding of a signal family $\mathcal{M}$ with conditioning δ if for all $x_1, x_2 \in \mathcal{M}$, we have the bi-Lipschitz property

$$(1-\delta)\|x_1 - x_2\|_2^2 \leq \|\Phi(x_1) - \Phi(x_2)\|_2^2 \leq (1+\delta)\|x_1 - x_2\|_2^2. \tag{1}$$

Notice that the conditioning δ indicates how well the measurement system maintains distances between points in $\mathcal{M}$. Small values of δ are good as they ensure that distances between signals are approximately preserved (i.e., $\|\Phi(x_1) - \Phi(x_2)\|_2^2 \approx \|x_1 - x_2\|_2^2$). If δ takes values close to one, then Φ is badly conditioned as there could be pairs of points x_1, x_2 such that $\|\Phi(x_1) - \Phi(x_2)\|_2^2 \approx 0$. Thus, measurement systems Φ that provide a stable embedding of $\mathcal{M}$ with a small conditioning δ will ensure that the representation $\Phi(x)$ for any signal $x \in \mathcal{M}$ is robust to perturbations by noise.

We can rewrite the stable embedding statement (1) in a different way that will provide an useful insight on the conditioning. Suppose Φ is linear and thus can be represented by a $M \times N$ matrix. Then Φ is a stable embedding of $\mathcal{M}$ with conditioning δ if for all $u \in U(\mathcal{M} - \mathcal{M})$,

$$\left| \|\Phi u\|_2^2 - 1 \right| \leq \delta,$$

where we recall that $U(\mathcal{M} - \mathcal{M}) := \left\{ \frac{x_1 - x_2}{\|x_1 - x_2\|_2} \mid x_1 \neq x_2 \in \mathcal{M} \right\}$. Equivalently, we can

write

$$\sup_{u \in U(\mathcal{M}-\mathcal{M})} \left| \|\Phi u\|_2^2 - 1 \right| \leq \delta. \tag{2}$$

If Φ is random matrix, then $\left| \|\Phi u\|_2^2 - 1 \right|$ is a *random process* dependent on u. Thus, (2) elucidates the fact that showing Φ provides a stable embedding with a certain conditioning δ is equivalent to showing that the *supremum* of the random process is upper bounded by δ. We will also find the following fact that relates a squared conditioning to a non-squared conditioning useful. For any vector y (e.g., $y = \Phi u$) such that $|\|y\|_2 - 1| < 1$,

$$|\|y\|_2 - 1| \leq \left| \|y\|_2^2 - 1 \right| \leq 3 \, |\|y\|_2 - 1| \, . \tag{3}$$

This fact can be proved as follows. For the left inequality, observe that $|\|y\|_2^2 - 1| = |(\|y\|_2 - 1)(\|y\|_2 + 1)| \geq |\|y\|_2 - 1|$, where we used the fact that $\|y\|_2 \geq 0$. For the right inequality, we have $|\|y\|_2 - 1| \geq |\|y\|_2 - 1|^2 = |(\|y\|_2^2 - 1) - 2(\|y\|_2 - 1)| \geq |(\|y\|_2^2 - 1)| - 2\,|(\|y\|_2 - 1)|$, where we used the reverse triangle inequality.

Measurement systems that exploit the low dimensional structure of a signal family can usually guarantee embeddings or stable embeddings of the family with a number of measurements M far below the ambient dimension N. This remarkable fact will be the common theme in the next sections.

2.3 The Johnson-Lindenstrauss Lemma

The use of a linear, compressive (measurement) operator Φ on a high-dimensional, *finite* signal family $\mathcal{M} \subset \mathbb{R}^N$ for the purpose of dimensionality reduction is well-studied in the computer science community. Many different types of operators Φ have been shown to support this type of dimensionality reduction, and the guarantees provided by Φ are collectively known as the Johnson-Lindenstrauss (JL) lemma. The JL lemma in its original form [80] states that there exists a Lipschitz mapping $\Phi : \mathbb{R}^N \to \mathbb{R}^M$ with $M \gtrsim \log(|\mathcal{M}|)$ such that Φ is a stable embedding of $\mathcal{M}$. Following this

discovery, constructive versions of the JL lemma using randomized linear operators have appeared in the literature.

The easiest construction of such randomized operators Φ are those that satisfy a *uniform* concentration of measure inequality [87], meaning that for any x belonging to the finite signal family $\mathcal{M}$,

$$\mathbb{P}\left\{\left|\|\Phi x\|_2^2 - \|x\|_2^2\right| > \delta\|x\|_2^2\right\} \leq 2e^{-Mc_0(\delta)}, \tag{4}$$

where $c_0(\delta)$ is some constant that depends only on δ (usually on the order of δ^2) and $\mathbb{E}\left\{\|\Phi x\|_2^2\right\} = \|x\|_2^2$. In words, an operator that satisfies (4) concentrates the energy of any given signal sharply around its mean with high probability. Moreover, for a fixed probability, the sharpness of the concentration (i.e., how small δ can be) correlates with the number of rows of Φ as can be seen by the exponential dependence on M by the probability bound in (4). Random matrices populated with i.i.d. subgaussian entries are known to satisfy the uniform concentration result of the form (4) [50, 96]. The JL lemma for any random operator Φ satisfying the concentration of measure bound (4) is given as follows:

Lemma 2.3.1 (Johnson-Lindenstrauss Lemma)**.** *[1, 42] Let $\mathcal{M} \subset \mathbb{R}^N$ be a finite signal family. Fix a conditioning $0 < \delta < 1$ and a failure probability $0 < \rho < 1$. Suppose $\Phi \in \mathbb{R}^{M\times N}$ is a random matrix satisfying the concentration of measure bound* (4) *for all $x \in \mathcal{M}$. If*

$$M \geq \frac{2}{c_0(\delta)} \log\left(\sqrt{\frac{2}{\rho}}|\mathcal{M}|\right), \tag{5}$$

then with probability exceeding $1-\rho$, Φ is a stable embedding of $\mathcal{M}$ with conditioning δ.

Similarly to the original JL lemma, Lemma 2.3.1 says that if we take a number of measurements M proportional to $\log(|\mathcal{M}|)$, then with high probability, Φ is a stable embedding of $\mathcal{M}$. We remark that the bound (5) has been shown to be optimal [4],

i.e., we cannot get a stable embedding of $\mathcal{M}$ with any less measurements (up to a constant multiple). The proof of Lemma 2.3.1 is straightforward.

Proof. Recall from (2) that a random matrix Φ is a stable embedding of a (finite) set $\mathcal{M}$ with conditioning δ if

$$\max_{u \in U(\mathcal{M}-\mathcal{M})} \left| \|\Phi u\|_2^2 - 1 \right| \leq \delta.$$

The goal is to calculate the failure probability of the above event, i.e.,

$$\mathbb{P}\left\{ \max_{u \in U(\mathcal{M}-\mathcal{M})} \left| \|\Phi u\|_2^2 - 1 \right| > \delta \right\}.$$

By the union bound and applying the concentration of measure inequality (4), we have

$$\begin{aligned} \mathbb{P}\left\{ \max_{u \in U(\mathcal{M}-\mathcal{M})} \left| \|\Phi u\|_2^2 - 1 \right| > \delta \right\} &\leq |U(\mathcal{M}-\mathcal{M})| \max_{u \in U(\mathcal{M}-\mathcal{M})} \mathbb{P}\left\{ \left| \|\Phi u\|_2^2 - 1 \right| > \delta \right\} \\ &\leq 2|U(\mathcal{M}-\mathcal{M})| e^{-Mc_0(\delta)}. \end{aligned}$$

Bounding the above by the chosen failure probability ρ and using the inequality $|U(\mathcal{M}-\mathcal{M})| \leq |\mathcal{M}|^2$ completes the proof. □

The application of the JL lemma to a finite point cloud is a useful pre-processing step for many database processing problems as it greatly improves processing speed while retaining pairwise distances between the points. One example of such a problem is the approximate nearest neighbor (ANN) problem [76] where given a finite set $\mathcal{M}$ and a query point $x \in \mathcal{M}$, the aim is to find a point $s \in \mathcal{M}$ (called the α-approximate nearest neighbor of x) such that for all $s' \in \mathcal{M}$, we have $\|x - s\|_2 \leq (1+\alpha)\|x - s'\|_2$ for a fixed tolerance α. Since projecting $\mathcal{M}$ into the lower dimensional space $\mathbb{R}^M$ by a operator Φ satisfying JL does not change distances between the points by much, we would not expect the performance of the ANN task to fare much worse in the measurement space than in the ambient space. More importantly, performing this

task in the measurement space greatly reduces the computational burden to far below what is required in the high-dimensional ambient space.

In addition, the JL lemma also has applications to compressive signal processing [43, 44, 57], where the goal is to perform various signal processing tasks (e.g., signal detection and estimation) in the compressed or measurement domain without first reconstructing the signal. We shall also see later that the JL lemma can be used to prove the stable embeddings of other signal families.

We remark that randomized operators satisfying the concentration of measure inequality (4) are typically unstructured (i.e., comprised of i.i.d. random entries). This lack of structure entails inefficiencies in the measurement process (i.e., it could be slow) and large memory storage requirements. To combat this, JL lemmas involving randomized operators that have fast transforms [3] and/or low storage requirements [2] have recently been introduced in the literature. These fast and efficient operators have been generalized in the paper [84], where the authors showed that all operators satisfying the Restricted Isometry Property[3] can be used for the JL lemma after randomizing their columns signs (see Theorem 2.4.3 below).

2.4 Compressed Sensing

Compressed Sensing (CS) [25, 53] is the study of compressive measurement systems of sparse signals and the design of computationally tractable algorithms to recover these signals from their underdetermined measurements. Because many signals of interest are sparse (in a certain basis), applications of CS have seen an explosive growth (e.g., the single pixel camera [56], MRI [91], radar systems [14], novel channel estimation schemes [71], and novel data acquisition systems [133]).

The success of CS depends strongly on signals being sparse or nearly sparse in a certain basis Ψ. To fix notations, we say that a signal $x \in \mathbb{C}^N$ is K-sparse if

[3]The Restricted Isometry Property will be described in Section 2.4.

$\|x\|_0 \leq K$,[4] and the value K is called the sparsity of the signal x. We say that a signal is nearly sparse if it is well represented by its (K) largest coefficients. We also say that a signal $x \in \mathbb{C}^N$ is K-sparse in an orthonormal basis $\Psi \in \mathbb{C}^{N \times N}$ if $\|\Psi^H x\|_0 \leq K$. Finally, denote the family of all K-sparse signals as $\Sigma_K \subset \mathbb{C}^N$, i.e.,

$$\Sigma_K := \{x \in \mathbb{C}^N \mid \|x\|_0 \leq K\}.$$

Many natural signals are sparse or nearly sparse in a certain basis (e.g., images are sparse in a wavelet basis [94]), and many signal processing algorithms are already taking advantage of this fact. Indeed, transform coders such as JPEG2000 exploit the near sparsity of images by first capturing a high-resolution image $x \in \mathbb{R}^N$ of a scene (say taken by a high resolution camera) then keeping only the K most significant coefficients in the wavelet basis (thus compressing the image). But this is a wasteful process since we have to first capture all these bits of the high-resolution image only to discard most of them in the compression phase. The novelty of CS is in introducing measurement systems that capture signals directly in the compressed form.

2.4.1 RIP and Sparse Signal Recovery

One important idea developed by CS is the notion of the Restricted Isometry Property (RIP) of a measurement operator.

Definition 2.4.1. *[25, 26] We say that a linear operator $\Phi \in \mathbb{C}^{M \times N}$ satisfies the RIP of* order K *with* conditioning δ*, or RIP-(K, δ) in short, if for all $x \in \Sigma_K$,*

$$(1 - \delta)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta)\|x\|_2^2. \tag{6}$$

We see immediately that the RIP of a measurement operator is closely related to the notion of stable embeddings (1) seen in Section 2.2. Indeed, if an operator Φ satisfies RIP-$(2K, \delta)$, then it is a stable embedding of Σ_K. As will be discussed in

[4] We use complex signals here to accommodate signals that are sparse in the Fourier basis (which are typically complex).

more details in Sections 2.4.2 and 2.4.3, measurement operators that satisfy the RIP typically have a number of rows M that scale linearly with the sparsity K (and only logarithmically with N) and M can be much less than the ambient dimension N. We remark that the RIP can be seen as an extension of the JL lemma from the stable embedding of a finite set to that of an infinite family of sparse signals.

The RIP is important in CS because it provides a sufficient condition for the robust recoverability of sparse (and nearly sparse) vectors from their noisy measurements.

Theorem 2.4.1 (Uniform CS Recovery with RIP)**.** *[111] Assume a matrix Φ satisfies RIP-$(2K, \delta)$ with $\delta < 0.4651$. Let $x \in \mathbb{C}^N$ be any vector and suppose we acquire the noisy measurements $y = \Phi x + e$ with $\|e\|_2 \leq \eta$. Let $\widehat{x}$ be the unique solution of:*

$$\min_z \|z\|_1 \quad \text{subject to} \quad \|\Phi z - y\|_2 \leq \eta. \tag{7}$$

Then

$$\|x - \widehat{x}\|_2 \leq C_1 \eta + C_2 \frac{\sigma_K(x)_1}{\sqrt{K}}, \tag{8}$$

where $\sigma_K(x)_1 := \inf_{z \in \Sigma_K} \|x - z\|_1$ is the error (measured in the ℓ_1-norm) of the best K-term approximation of x, and C_1, C_2 depend only on δ.

The convex optimization program (7) is known as Basis Pursuit De-Noising (BPDN) [37], and many algorithms exist that solve it (e.g., [17, 60, 149]). If a signal x is exactly K-sparse and the measurement process is noiseless ($\eta = 0$), then $\|x - \widehat{x}\|_2 = 0$ in (8) which implies a perfect recovery of x. For a general signal x whose measurements y are corrupted by bounded noise ($0 < \eta < \infty$), solving the BPDN program (7) guarantees an output $\widehat{x}$ whose distance from x is bounded both by the measurement noise level (η) and by the distance from x to its best K-term approximation ($\sigma_K(x)_1$). When x is nearly sparse, $\sigma_K(x)_1$ is small and thus BPDN recovers a good approximation of x. Most importantly, this theorem highlights the surprising result of CS; given that a measurement operator $\Phi \in \mathbb{R}^{M \times N}$ satisfies RIP-$(2K, \delta)$ for M much less than N, any sparse signal can be recovered from their undersampled measurements by solving (7).

Greedy iterative algorithms such as CoSAMP [101], IHT [20], and OMP [132] can also used (instead of BPDN) to recover sparse signals from their measurements. As a testament to the strength of stable embeddings, the RIP can also provide recoverability guarantees for OMP [46] and CoSAMP [101].

2.4.2 RIP for Subgaussian Random Matrices

Because of their amenability to analysis, unstructured, random matrices (i.e., matrices with i.i.d. random entries) have initially drawn much interest as measurement systems for CS. Of particular interest are random $M \times N$ matrices with i.i.d. zero-mean, subgaussian random variables with variance $\frac{1}{M}$ (from here on called *subgaussian random matrices*). First, recall from Section 2.3 that all subgaussian random matrices have a uniform concentration of measure inequality as expressed in (4). From this property, we can show that all subgaussian random matrices Φ satisfy the RIP-(K, δ) with high probability whenever M scales linearly with K and logarithmically with N:

Theorem 2.4.2. *[13] Suppose $\Phi \in \mathbb{R}^{M \times N}$ is a subgaussian random matrix satisfying the uniform concentration of measure inequality* (4) *and let $\Psi \in \mathbb{R}^{N \times N}$ be any basis of $\mathbb{R}^N$. Let δ and ρ be the predetermined RIP conditioning and failure probability respectively. If*

$$M \gtrsim \frac{1}{c_0(\delta)} \left[K \log\left(\frac{N}{K\delta}\right) + \log\left(\frac{1}{\rho}\right) \right],$$

then $\Phi\Psi$ satisfies RIP-(K, δ) with probability greater than $1 - \rho$.

The measurement rate given by this theorem (i.e., the number of measurements M needed to make a particular RIP guarantee) is optimal. This fact arises from a beautiful connection of CS with geometry, in particular the works of Garnaev, Gluskin, and Kashin on n-widths [65, 82] (see also [13]). Second, this theorem also says that subgaussian random matrices are *universal* measurement matrices where the measurement rate M of Φ does not depend on the sparsity basis.

The proof of this theorem uses simple covering and counting arguments together with the union bounds [13]. As some of our results use the same techniques, we shall provide a proof sketch of Theorem 2.4.2 here. First, $\Phi\Psi$ satisfies RIP-(K,δ) if and only if $\sup_{u\in U(\Sigma_K)} \left|\|\Phi\Psi u\|_2^2 - 1\right| \leq \delta$. We then wish to estimate the failure probability

$$\mathbb{P}\left\{ \sup_{u\in U(\Sigma_K)} \left|\|\Phi\Psi u\|_2^2 - 1\right| > \delta \right\}.$$

Now, the set $U(\Sigma_K)$ has an infinite number of elements, and this prevents a simple application of the union bound as done in the proof of the JL lemma (Lemma 2.3.1). To proceed, first notice that $U(\Sigma_K)$ is a union of K-dimensional unit ℓ_2-balls of $\mathbb{R}^N$ that makes up the space of all unit-norm K-sparse vectors. We can *cover* each of these K-dimensional balls with a finite covering set (see [136, Lemma 2]). The union bound is then taken over all these finite covering sets. Following this, simple geometric and algebraic arguments are used to extend the failure probability from the covering sets to the whole of $U(\Sigma_K)$.

2.4.3 RIP for Structured Random Matrices

Recently, the CS community has turned to investigating structured measurement systems because unstructured systems may be impractical due to memory constraints, computational costs, or limitations in the data acquisition architecture. One example of a structured random matrix is a subsampled Fourier matrix formed by picking M rows of a $N \times N$ Discrete Fourier Transform (DFT) matrix uniformly at random. Rudelson and Vershynin [118] showed that when the number of rows satisfies $M \gtrsim \frac{K}{\delta^2}\log^4(N)$, then subsampled Fourier matrices satisfy RIP-(K,δ) with high probability. Because Fourier transforms are fast to compute, measurement systems represented by such matrices are computationally efficient. Moreover, subsampled Fourier matrices represent realistic MRI measurement systems [91]. This RIP result is generalized in [111] (and later improved in [6]) to include all subsampled and bounded orthonormal systems.

Another family of structured random matrices of interest are subsampled circulant or Toeplitz matrices representing the convolution of a random probe with an input signal. These measurement matrices are of interest as they represent radar systems [14] and novel channel estimation schemes [11, 71]. By supposing that the Fourier transform of the probe has i.i.d. random entries, Romberg showed that subsampled circulant measurement systems satisfy RIP-(K, δ) whenever the number of rows M scales linearly with the sparsity K and poly-logarithmically with the ambient dimension N (i.e., $\log^{\alpha}(N)$ with $N > 1$) [113, 114]. If instead the probe itself (not its Fourier transform) has i.i.d. random entries, Krahmer et al. [83] (improving on the work in [112]) showed these measurement systems satisfy RIP-(K, δ) with high probability whenever the number of rows $M \gtrsim \frac{K}{\delta^2} \log^4(N)$.

As a final example of a structured measurement system, Tropp et al. [133] showed that a random demodulator matrix representing a practical CS measurement system (that potentially replaces a traditional A/D converter) satisfies RIP-(K, δ) with high probability whenever $M \gtrsim \frac{K}{\delta^2} \log^6(N)$.

Structured random matrices are typically not universal and therefore do not satisfy the RIP for signals sparse in all bases. For example, the subsampled Fourier matrix achieves an optimal measurement rate when measuring canonical sparse signals (i.e. $x \in \Sigma_K$) and will require more measurements when used for signals sparse in other bases [111]. Indeed, the measurement rates for structured random matrices (to satisfy RIP) typically scale with an *incoherence* property that depends on the sparsity basis [29]. This incoherence property will be made clear in our work on block diagonal matrices in Chapter 3 and the short-term memory of Echo State Networks in Chapter 6.

The proofs of the RIP for structured random matrices usually require sophisticated probabilistic tools to bound the extrema of random processes. We will provide a short outline of these proof techniques here to contrast against the simple arguments used

to prove Theorem 2.4.2. We will again be using them in the proofs of some of our own theorems.

The first technique (with a slight abuse of notation) starts by defining the Restricted Isometry Constant (RIC) of order K of a measurement operator $\Phi \in \mathbb{C}^{M\times N}$ as

$$\delta_K = \delta_K(\Phi) := \sup_{u\in U(\Sigma_K)} \left| \|\Phi u\|_2^2 - 1 \right| = \sup_{u\in U(\Sigma_K)} \left| u^H(\Phi^H\Phi - I)u \right|.$$

This technique is used particularly for matrices Φ that have independent (in the probabilistic sense) rows (e.g., a subsampled Fourier matrix). With this definition, Φ satisfies RIP-(K,δ) if and only if $\delta_K \leq \delta$. As always, we will want to show that the probability that δ_K exceeds δ (i.e., $\mathbb{P}\{\delta_K > \delta\}$) is small. Denote $\phi_1^H, \cdots, \phi_M^H \in \mathbb{C}^N$ as the independent rows of the matrix Φ. Then, $\Phi^H\Phi = \sum_{m=1}^M \phi_m\phi_m^H$, where we see that the Grammian matrix $\Phi^H\Phi$ can be written as an outer-product of *rank-one* matrices. Next, we define a norm on Hermitian operators called the RIP-norm as

$$|||B||| := \sup_{u\in U(\Sigma_K)} \left| u^H B u \right|.$$

We can check that this definition fulfills all the conditions of a norm, and we see that the RIC of an operator Φ is simply the RIP-norm of the Hermitian operator $\Phi^H\Phi - I$ (i.e., $\delta_K = \left|\left|\left|\Phi^H\Phi - I\right|\right|\right|$). As the matrices under consideration usually have rows that satisfy $\mathbb{E}\left\{\phi_m\phi_m^H\right\} = \frac{1}{M}I$ for all $m = 1, \cdots, M$, we have $\mathbb{E}\left\{\Phi^H\Phi\right\} = I$, and thus,

$$\delta_K = \left|\left|\left|\Phi^H\Phi - \mathbb{E}\left\{\Phi^H\Phi\right\}\right|\right|\right| = \left|\left|\left| \sum_{m=1}^M \phi_m\phi_m^H - \mathbb{E}\left\{\phi_m\phi_m^H\right\} \right|\right|\right|.$$

Then, the crux of this first proof method hinges on Lemma A.1.2 in Section A.1, which gives the tail probability of a random variable Z once we know its moments $(\mathbb{E}\{|Z|^p\})^{1/p}$. Therefore, the goal will be to calculate the moments $(\mathbb{E}\{\delta_K^p\})^{1/p}$. To do so, we first apply the symmetrization trick as shown in Lemma A.1.1 to obtain

$$(\mathbb{E}\{\delta_K^p\})^{1/p} \leq 2\left(\mathbb{E}\left\{ \left|\left|\left| \sum_{m=1}^M \xi_m\phi_m\phi_m^H \right|\right|\right|^p \right\} \right)^{1/p}, \tag{9}$$

where $\{\xi_m\}$ is a Rademacher sequence independent from $\{\phi_m\}$ and the expectation on the right is over *both* these sequences. To bound the right-hand-side of (9), we use the following lemma (aptly called the "Uniform Law of Large Numbers" by Rudelson and Vershynin in [118]):

Lemma 2.4.1 (Uniform Law of Large Numbers). *[111, 118] For $m = 1, \cdots, M$, let $\phi_m \in \mathbb{C}^N$ be (fixed) vectors such that $\|\phi_m\|_\infty \leq L < \infty$, and let $\{\xi_m\}$ be a Rademacher sequence. Then, for $p \geq 2$,*

$$\begin{aligned} &\left(\mathbb{E} \left\{ \left\| \sum_{m=1}^{M} \xi_m \phi_m \phi_m^H \right\|^p \right\} \right)^{1/p} \\ \leq \quad & \beta^{1/p} C \sqrt{p} L \sqrt{K} \log(100K) \sqrt{\log(4N)\log(10M)} \sqrt{\left\| \sum_{m=1}^{M} \phi_m \phi_m^H \right\|}, \end{aligned}$$

where $C \approx 67.97$ and $\beta = 6.028$, and the expectation is over the Rademacher sequence $\{\xi_m\}$.

A second proof techniques for the RIP of structured random matrices uses Dudley's inequality (which incidentally is also used in the proof of the Uniform Law of Large Numbers, Lemma 2.4.1). This method is used for example in [83, 112] for subsampled circulant matrices where the RIC can be written as

$$\delta_K := \sup_{u \in U(\Sigma_K)} \langle \xi, Z_u \xi \rangle, \tag{10}$$

where $\xi \in \mathbb{R}^N$ is a Rademacher sequence representing the probe, and $Z_u \in \mathbb{C}^{N \times N}$ is a hollow (meaning that the diagonal entries of Z_u are all zeros) Hermitian matrix that depends on $u \in U(\Sigma_K)$. Written this way, the RIC δ_K is the supremum of a Rademacher chaos random process (dependent on $u \in U(\Sigma_K)$), and thus, a version of Dudley's inequality for Rademacher chaos processes can be used to calculate its expectation:[5]

[5] We note that a sharper bound for the expectation can be found in [83] (which is used in our proof of Theorem 3.2.2 in Chapter 3) but not shown here as it requires introducing additional notations.

Lemma 2.4.2 (Dudley's Inequality for Rademacher Chaos)**.** *[112] Suppose that for any $u \in \mathcal{S}$, $Z_u \in \mathbb{C}^{N\times N}$ is a hollow Hermitian matrix. Fix a point $u_0 \in \mathcal{S}$ and let $\xi \in \mathbb{R}^N$ be a Rademacher sequence. Then, there exists a universal constant C such that*

$$\begin{aligned} & \mathbb{E}\left\{\sup_{u\in\mathcal{S}} |\langle \xi, (Z_u - Z_{u_0})\xi\rangle|\right\} \\ \leq \;& C\max\left\{\int_0^\infty \log\left(|\mathcal{C}\left(\mathcal{S}, d_1, \eta\right)|\right) d\eta, \int_0^\infty \sqrt{\log\left(|\mathcal{C}\left(\mathcal{S}, d_2, \eta\right)|\right)} d\eta\right\}, \end{aligned}$$

where $d_1(u_1, u_2) := \|Z_{u_1} - Z_{u_2}\|_2$ and $d_2(u_1, u_2) := \|Z_{u_1} - Z_{u_2}\|_F$ are called the subexponential *and* subgaussian *distances respectively.*

Since $0 \in U\left(\Sigma_K\right)$, we will set $u_0 = 0$ in Lemma 2.4.2. The main difficulty in applying this lemma is in calculating the covering numbers $|\mathcal{C}\left(U\left(\Sigma_K\right), d_i, \eta\right)|$ for $i = 1, 2$. Once we have estimated the expectation $\mathbb{E}\left\{\delta_K\right\}$ with Lemma 2.4.2, various methods exist to show that δ_K does not deviate much from its expectation (e.g., see Lemma F.3.2 in Appendix F.3).

We remark that not all matrices that satisfy the RIP are random in nature. For example, in [49] DeVore presented a deterministic construction of a measurement matrix that satisfies RIP-(K, δ) whenever $M \gtrsim \frac{K^2}{\delta^2}\log^2(N)$. Despite the additional number of required measurements, deterministic matrices can be of interest to the CS community as it is an NP-hard problem to verify whether a randomly constructed matrix satisfies the RIP [130].

We also remark that there is a close connection between the RIP and the JL lemma. This is made clear by the following theorem:

Theorem 2.4.3. *[84] Fix $0 < \rho, \epsilon < 1$ and suppose there is a finite set of points $E \subset \mathbb{R}^N$. Also suppose we have a matrix $\Phi \in \mathbb{R}^{M\times N}$ satisfying the RIP of order $k \geq 40\log\left(\frac{4|E|}{\rho}\right)$ and conditioning $\delta \leq \frac{\epsilon}{4}$. Let $\xi \in \mathbb{R}^N$ be a Rademacher sequence, construct the diagonal Rademacher matrix $D_\xi := \mathrm{diag}(\xi)$, and define $\widehat{\Phi} := \Phi D_\xi$.*

Then with probability exceeding $1-\rho$, *we have for all* $x \in E$,

$$(1-\epsilon)\|x\|_2^2 \leq \|\widehat{\Phi}x\|_2^2 \leq (1+\epsilon)\|x\|_2^2.$$

In words, any operator satisfying the RIP can be used to approximately preserve the norms of the signals in a given finite point cloud when the signs of the columns of the operator are randomly chosen. We remark that if the finite point cloud E is the set of all differences between points in another finite set $\mathcal{M} \subset \mathbb{R}^N$, then a matrix $\Phi \in \mathbb{R}^{M \times N}$ satisfying the RIP of order $S \geq 40 \log\left(\frac{4|\mathcal{M}|^2}{\rho}\right)$ (and conditioning $\delta \leq \frac{\epsilon}{4}$) in Theorem 2.4.3 can provide a stable embedding of $\mathcal{M}$ with high probability when the column signs of Φ are randomized.

2.5 Embeddings of Manifolds

The sparsity and low-rank signal models that have gained significant attention in the signal processing community do not apply well to all signal families. Nonetheless, many high-dimensional signals can generally be modeled as lying on low-dimensional submanifolds embedded in Euclidean space.

A common example of a manifold-modeled signal family is an image appearance manifold (IAM) [137]. Each point in an IAM is an image of a (3-dimensional) object, and a parameter $\theta \in \Theta$ controls the way an object appears in the image. The image of the object produced by a certain $\theta \in \Theta$ can be represented by a function $f_\theta : \mathbb{R}^2 \to \mathbb{R}$ and the IAM $\mathcal{F}$ is simply the collection of all images (i.e., $\mathcal{F} := \{f_\theta \mid \theta \in \Theta\}$). For example, the object could be a (simulated) bust of a person's head and the set Θ could consist of all different orientation of the bust and all different orientations of the lighting (as shown in Figure 1 of Chapter 1). When we consider a finite-resolution representation of the images, a vector $x_\theta \in \mathbb{R}^N$ can be used to represent each image f_θ. The corresponding set of all vectors x_θ will also lie on a manifold $\mathcal{M}$, i.e., $\mathcal{M} := \{x_\theta \mid \theta \in \Theta\} \subset \mathbb{R}^N$.[6]

[6]By consider a finite-resolution version of the images, we side-step the nowhere-differentiability

Apart from *parametric* signal families as shown by the IAM, manifold models can also be used for certain *non-parametric* signal families. These signal families include the set of images of hand-written digits [86] as shown in Figure 2 and the set of images of all human faces [134]. While it is hard to find a well-defined parameter that describes each non-parametric signal family, empirical evidence suggest that such signals do cluster near a low-dimensional manifold [74, 115, 129].

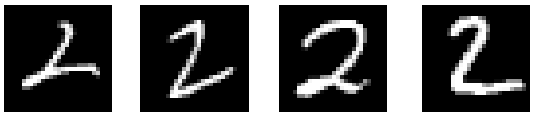


Figure 2: Examples of handwritten digits that lie close to a low-dimensional manifold in the ambient space [74, 86].

2.5.1 Some Characterizations of Riemannian Submanifolds

The analysis of measurement operators of manifold-modeled signals requires us to consider two additional characterizations of a manifold that will be useful for describing certain local and global properties of the manifold. The first is the *second fundamental form* of the manifold (as defined in [52]), which provides a bound on the worse case curvature of any unit speed geodesic path along the manifold. We assume that the second fundamental form is uniformly bounded by some number $\frac{1}{\tau}$, where this upper bound is related to the condition number (also denoted by $\frac{1}{\tau}$) as described in [102] and used in [15].

To describe the second fundamental form, we require the notion of the *standard connection* operator acting on vector fields of a submanifold $\mathcal{M} \subset \mathbb{R}^N$. Let $\nabla : \mathcal{X}(\mathcal{M}) \times \mathcal{X}(\mathcal{M}) \to \mathcal{X}(\mathbb{R}^N)$ denote the standard connection, where we write $\nabla_X Y$ to represent a resulting vector field formed from two other vector fields X, Y [52]. In our setting, the mapping ∇ generalizes the differentiation of a vector field along a curve on $\mathcal{M}$. More concretely, let $\gamma : [0,\ 1] \to \mathcal{M}$ be the parameterization of a curve on the

problem of IAM. For more details on this issue, see [54, 137].

manifold. Thus, $\gamma'(t) = \frac{d\gamma}{dt}(t)$ is an element of the tangent plane $\mathcal{T}_{\gamma(t)}\mathcal{M}$ and belongs to some vector field on $\mathcal{M}$. Let $v(t)$ denote a parameterization of another vector field along the curve γ (i.e., at every point $\gamma(t)$, we can associate a vector $v(t) \in \mathcal{T}_{\gamma(t)}\mathcal{M}$). For example, $v(t)$ could represent the velocity of a particle moving along γ. Then, the derivative of the vector field v along the curve γ (which is a vector field of the Euclidean space that not necessarily confined on $\mathcal{X}(\mathcal{M})$) is given by $v' = \frac{dv}{dt} := \nabla_{\gamma'} v$. If $v(t)$ represents velocity, then $\nabla_{\gamma'} v(t)$ represents the acceleration necessary to keep a particle long the path γ on the manifold and maintain its velocity.

Now, note that at every point $p \in \mathcal{M}$, the tangent plane of the *ambient space* at p can be split into two orthogonal components, namely $\mathcal{T}_p\mathcal{M}$ (which is the tangent space of the manifold $\mathcal{M}$ at point p) and its orthogonal component $(\mathcal{T}_p\mathcal{M})^\perp$. For any pairs of vector fields X, Y on the manifold $\mathcal{M}$, define $B(X,Y) := (\nabla_X Y)^\perp$ which is a vector field on the ambient space but lying solely on the orthogonal component $(\mathcal{T}_p\mathcal{M})^\perp$ at every point p. For a (velocity) vector field $v(t)$ along a curve $\gamma(t)$ on the manifold $\mathcal{M}$, $B(\gamma', v)$ is a (acceleration) vector field that is pointing normal to $\mathcal{M}$. $B(X,Y)$ can be shown to be a *symmetric, bilinear* form [52]. Thus, for every $p \in \mathcal{M}$ and for any vector η on $(\mathcal{T}_p\mathcal{M})^\perp$, $H_\eta(X,Y)_p := \langle B(X,Y)_p, \eta \rangle$ is also a symmetric, bilinear form, where $\langle \cdot, \cdot \rangle$ denote the standard inner product in Euclidean space. Consequently, we can find a $N \times N$ self-adjoint matrix S_η such that $H_\eta(X,Y)_p = \langle X_p, S_\eta Y_p \rangle$. Finally, the *second fundamental form* $\mathbb{I}_\eta(v,p)$ of the submanifold $\mathcal{M}$ at point p along the normal vector $\eta \in (\mathcal{T}_p\mathcal{M})^\perp$ for a tangent vector $v \in \mathcal{T}_p\mathcal{M}$ is given by

$$\mathbb{I}_\eta(v,p) := H_\eta(v,v)_p = \langle v, S_\eta v \rangle.$$

We say that the second fundamental form of $\mathcal{M}$ is *uniformly bounded* by $\frac{1}{\tau}$ if[7]

$$\sup_{v \in \mathcal{X}(\mathcal{M}), p \in \mathcal{M}, \eta \in (\mathcal{T}_p\mathcal{M})^\perp} |\mathbb{I}_\eta(v,p)| = \sup_{v \in \mathcal{X}(\mathcal{M}), p \in \mathcal{M}, \eta \in (\mathcal{T}_p\mathcal{M})^\perp} |\langle v, S_\eta v \rangle| \leq \frac{1}{\tau}.$$

[7]We bound the second fundamental form by a fraction $\frac{1}{\tau}$ to make explicit the correspondence to the condition number represented by the same fraction.

In words, this means that the spectral norm of the self-adjoint matrix S_η is bounded by $\frac{1}{\tau}$ (i.e., $\|S_\eta\|_2 \leq \frac{1}{\tau}$) for all vector fields v, for all points p, and for all normal vectors η in $(\mathcal{T}_p\mathcal{M})^\perp$.

The following lemma lists some consequences of the uniform boundedness of the second fundamental form on certain geometric properties of the manifold that will be useful to our analysis.

Lemma 2.5.1. *Suppose a submanifold $\mathcal{M} \in \mathbb{R}^N$ has second fundamental form uniformly bounded by $\frac{1}{\tau}$. Let $p, q \in \mathcal{M}$ be two distinct points. Then, we have the following three properties of the manifold:*

1. *(Curvature) If $\gamma(t)$ denotes a unit speed parameterization of the geodesic path joining p and q, then $\|\gamma''(t)\|_2 \leq \frac{1}{\tau}$. Moreover, denoting $\mu := d_{\mathcal{M}}(p, q)$, we have $q - p = \gamma(\mu) - \gamma(0) = \mu\gamma'(0) + R$ with $\|R\|_2 \leq \frac{\mu^2}{2\tau}$.*

2. *(Twisting of Tangent Spaces) Suppose $d_{\mathcal{M}}(p, q) \leq \tau$. Pick $u \in \mathcal{T}_p\mathcal{M}$, and let $v \in \mathcal{T}_q\mathcal{M}$ be the parallel transport[8] of u into $\mathcal{T}_q\mathcal{M}$. Then, $\langle u, v\rangle \geq 1 - \frac{d_{\mathcal{M}}(p,q)}{\tau}$.*

3. *(Self-avoidance) Suppose $\|p - q\|_2 \leq \frac{\tau}{2}$. Then, $\|p - q\|_2 \geq d_{\mathcal{M}}(p, q) - \frac{d_{\mathcal{M}}(p,q)^2}{2\tau}$. As a corollary, we also have $d_{\mathcal{M}}(p, q) \leq \tau - \tau\sqrt{1 - \frac{2\|p-q\|_2}{\tau}}$.*

Proof. The proof can be found in Appendix B.1. □

The first property says that the worst case curvature of any unit speed geodesic path along the manifold is bounded by $\frac{1}{\tau}$. The second property states that for small geodesic distances, the tangent spaces do not "twist" too much from one another. Thus, if we compare a tangent vector to its parallel counterpart in another nearby tangent space, the angle between them is small. The last property states that for

[8] *Suppose $\gamma(t)$ denotes a unit speed parameterization of the geodesic path joining p and q. By parallel transport [52], we mean a vector field $v(t)$ defined along $\gamma(t)$ such that $v(0) = u$, $v(\mu) = v$, $\|v(t)\|_2 = \|u\|_2$, and $\langle v(t), \gamma'(t)\rangle = \langle u, \gamma'(0)\rangle$, where the last two conditions mean that $v(t)$ maintains a constant length and angle with respect to the path $\gamma(t)$.*

points on the submanifold close together in Euclidean space, their geodesic and Euclidean distances do not differ much. Negating the statement, we see that two points with large geodesic distance cannot be arbitrarily close in Euclidean space.

A second useful quantity concerns the geodesic regularity of a manifold. Before going on to describe this quantity, let us introduce some pre-requisite terminology. For a Riemannian manifold $\mathcal{M}$, recall that $\mathcal{B}^{\mathcal{M}}_{d_{\mathcal{M}}}(x,\epsilon)$ is the geodesic ball centered at $x \in \mathcal{M}$ of radius ϵ (i.e., $\mathcal{B}^{\mathcal{M}}_{d_{\mathcal{M}}}(x,\epsilon) := \{p \in \mathcal{M} \mid d_{\mathcal{M}}(p,x) \leq \epsilon\}$). Similarly, recall that $\mathcal{B}^{K}_{2}(x,\epsilon)$ is the Euclidean ball of radius ϵ centered at $x \in \mathbb{R}^K$ (i.e., $\mathcal{B}^{K}_{2}(x,\epsilon) := \{p \in \mathbb{R}^K \mid \|p-x\|_2 \leq \epsilon\}$). Then, the geodesic regularity R is defined as follows:

Definition 2.5.1. *A K-dimensional Riemannian submanifold $\mathcal{M}$ of $\mathbb{R}^N$ has* geodesic regularity R *if for every $\epsilon \leq \epsilon_0$ and for every $x \in \mathcal{M}$,*

$$\operatorname{vol}\left(\mathcal{B}^{K}_{2}(x,\epsilon)\right) \leq R^K \operatorname{vol}\left(\mathcal{B}^{\mathcal{M}}_{d_{\mathcal{M}}}(x,\epsilon)\right).$$

We see that the geodesic regularity allows a uniform comparison of the geodesic and Euclidean balls of the same radius everywhere on the manifold. This comparison is related to a certain intrinsic curvature (in particular, the scalar curvature) of the manifold [69]. As in [15], we shall subsequently neglect the minor dependence of the geodesic regularity R on the maximum resolution ϵ_0.

The geodesic regularity R of a manifold $\mathcal{M}$ allows us to quantify the geodesic covering number of the manifold (i.e., how many geodesic balls of a certain radius are needed to cover the whole manifold). The following lemma gives an upper bound on the geodesic covering number of a manifold.

Lemma 2.5.2. *The $(\epsilon, d_{\mathcal{M}})$-covering number of a compact K-dimensional Riemannian submanifold $\mathcal{M} \subset \mathbb{R}^N$ is bounded by*

$$\left|\mathcal{C}\left(\mathcal{M}, d_{\mathcal{M}}, \epsilon\right)\right| \leq \frac{V}{\inf_{x \in \mathcal{M}} \operatorname{vol}\left(\mathcal{B}^{\mathcal{M}}_{d_{\mathcal{M}}}(x, \frac{\epsilon}{2})\right)},$$

where $V := \mathrm{vol}(\mathcal{M})$. *If* $\mathcal{M}$ *has geodesic regularity* R, *then*

$$|\mathcal{C}(\mathcal{M}, d_{\mathcal{M}}, \epsilon)| \leq \frac{\left(\frac{2R}{\sqrt{\pi}}\right)^K \left(\sqrt{K/2+1}\right)^K V}{\epsilon^K}. \tag{11}$$

Proof. The proof can be found in Appendix B.2. □

We remark that the definition of an equivalent geodesic covering regularity in [15] corresponds to $\frac{2R}{\sqrt{\pi}}$ appearing in (11).

2.5.2 Embeddings

Answering the question of whether a manifold can equivalently be defined with or without an extrinsic Euclidean space, Whitney [140] showed that a manifold $\mathcal{M}$ of dimension K can be smoothly embedded into Euclidean space $\mathbb{R}^M$ as long as $M \geq 2K$. Whitney's embedding theorem is later strengthened by Nash [100] for Riemannian manifolds. Nash showed that not only can there be an embedding (specifically a $\mathcal{C}^1$ embedding), the embedding can also be an isometry. This means that when appropriately mapped by the embedding operator, the Riemannian metric of the manifold corresponds to the induced metric from the Euclidean space. This isometric embedding ensures that geodesic distances (measured by the intrinsic and induced metric respectively) between points on the manifold are maintained in the image manifold. Thus, Nash's embedding theorem answered the question of whether it is equivalent to define the Riemannian metric intrinsically or extrinsically with an Euclidean metric.

In this thesis, we consider signal families modeled by compact low-dimensional Riemannian submanifolds embedded in high-dimensional Euclidean space. In this context, Whitney's and Nash's embedding results have a different applicability. These embedding results tell us that if we have a K-dimensional Riemannian submanifold $\mathcal{M}$ initially embedded in high-dimensional Euclidean space $\mathbb{R}^N$ with $K \ll N$, then there exists an (isometric) embedding $\Phi : \mathbb{R}^N \to \mathbb{R}^M$ with $M \geq 2K$ (and $M \ll N$)

that maps $\mathcal{M} \subset \mathbb{R}^N$ into a lower-dimensional Euclidean space $\mathbb{R}^M$. In particular, Φ can represent a compressive measurement system whose outputs can be used to distinguish between different signals from $\mathcal{M}$ despite a reduction in dimensionality. Although it proves the existence of embeddings that preserve geodesic distances, Nash's embedding result still suffers from two drawbacks that we will address in this thesis. First, Nash's result is not constructive and thus does not provide us with any practical embedding operator. Second, the preservation of just geodesic distances on a submanifold does not imply that Euclidean distances are preserved as well. This means that manifold characteristics such as curvature may not be preserved in the embedding space, and thus the geometry of the embedded submanifold may be very different from that of the ambient submanifold.

2.5.3 Stable Embeddings

Having an embedding without consideration of the eventual distances between points on the manifold $\mathcal{M}$ in the measurement space may not be enough in noisy scenarios. Recall that points that were initially far away may be mapped close together by the embedding and they may become indistinguishable under noise.

It was Baraniuk and Wakin [15] who extended these embedding results to show stable embeddings of manifolds. Concretely, their result is as follows:

Theorem 2.5.1. *[15] Let $\mathcal{M}$ be a compact K-dimensional Riemannian submanifold of $\mathbb{R}^N$ with second fundamental form uniformly bounded by $\frac{1}{\tau}$, volume V and geodesic regularity R. Fix $0 < \delta < 1$ and $0 < \rho < 1$. Let Φ be a random orthoprojector from $\mathbb{R}^N$ to $\mathbb{R}^M$ with*

$$M \gtrsim \frac{1}{\delta^2}\left[K \log\left(\frac{RN}{\tau\delta}\right) + \log\left(\frac{V}{\rho}\right)\right] \log\left(\frac{N}{K}\right).$$

Then, with probability at least $1-\rho$, $\sqrt{\frac{N}{M}}\Phi$ is a stable embedding of $\mathcal{M}$ with conditioning δ.

To be precise, a random orthoprojector from $\mathbb{R}^N$ to $\mathbb{R}^M$ can be expressed as a matrix with orthonormal rows constructed by first generating M length-N vectors with i.i.d. zero-mean Gaussian random entries and then applying a Gram-Schmidt algorithm. This theorem shows that K-dimensional Riemannian submanifolds $\mathcal{M}$ that are embedded in an (high-dimensional) Euclidean space $\mathbb{R}^N$ can be stably embedded with high probability by random and non-adaptive orthoprojections into $\mathbb{R}^M$. In correspondence to the RIP results in CS, the measurement rate M simply needs only to scale linearly with K, logarithmically with N, and logarithmically with properties of the manifold including the second fundamental form, volume, and geodesic regularity. Clarkson [39] later improved on the required number of measurements M by removing the dependence on N and certain worst case properties of the manifold.[9]

A proof sketch of Theorem 2.5.1 is given below. As in the proof of the RIP of subgaussian random matrices (Theorem 2.4.2), $\widehat{\Phi} := \sqrt{\frac{N}{M}}\Phi$ is a stable embedding of $\mathcal{M}$ with conditioning δ if and only if $\sup_{u \in U(\mathcal{M}-\mathcal{M})} \left| \left\| \widehat{\Phi} u \right\|_2^2 - 1 \right| \leq \delta$ where we recall that $U(\mathcal{M}-\mathcal{M}) = \left\{ u = \frac{x-y}{\|x-y\|_2} \mid x \neq y \in \mathcal{M} \right\}$. We then wish to estimate the failure probability

$$\mathbb{P}\left\{ \sup_{u \in U(\mathcal{M}-\mathcal{M})} \left| \left\| \widehat{\Phi} u \right\|_2^2 - 1 \right| > \delta \right\}.$$

First, random orthoprojectors $\widehat{\Phi}$ satisfy a uniform concentration of measure (4) (see Example 25.2 in [136]). This means that by carefully choose a finite covering set of $U(\mathcal{M}-\mathcal{M})$, we can apply the union bound on the covering set with the concentration of measure bound of $\widehat{\Phi}$. Then, geometric arguments that utilize the manifold characterizations can be used to extend the stable embedding from the finite covering set to the whole of $U(\mathcal{M}-\mathcal{M})$.

Stable embeddings of manifolds by compressive operators are valuable because

[9] Clarkson removed the dependence on the ambient dimension N by means of chaining arguments that appear in proofs of Dudley's inequality. Because our results follow more closely that of [15] instead of [39], we will not elaborate on the latter here.

they ensure that key properties of the manifold are retained in the low-dimensional measurement space. First, similarly to how the RIP ensures recoverability of sparse signals, the stable embedding of a manifold leads to guarantees on the recoverability of signals residing on the manifold from underdetermined measurements. The recovery algorithms used are typically iterative projection algorithms, and one example of such an algorithm is the Manifold Iterative Projection (MIP) algorithm proposed by Shah and Chandrasekaran [122]. In correspondence to recoverability guarantees in CS (Theorem 2.4.1), the recoverability guarantee for MIP goes as follows:

Theorem 2.5.2. *[122] Suppose Φ is a stable embedding of a K-dimensional Riemannian submanifold $\mathcal{M}$ with conditioning $\delta < \frac{1}{3}$, and suppose we observe the measurements $y = \Phi x$. Then, the MIP algorithm converges, i.e., $\widehat{x}_T \to x$ as $T \to \infty$. Moreover, denoting $\widehat{x}_T \in \mathcal{M}$ as the output of the T-th iterate of the MIP, we have $\|\Phi\widehat{x}_T - y\|_2^2 \leq \epsilon$ and $\|\widehat{x}_T - x\|_2^2 \leq \frac{\epsilon}{1-\delta}$ in $T = \left\lceil \frac{1}{\log\left(\frac{1-\delta}{2\delta}\right)} \log\left(\frac{\|y\|_2^2}{2\epsilon}\right) \right\rceil$ iterations.*

As Theorem 2.5.2 indicates, a stable embedding by the measurement operator Φ ensures that the MIP algorithm converges. Moreover, the algorithm converges at a rate proportional to $\log\left(\frac{1}{\epsilon}\right)$ for any error tolerance ϵ.

Second, the stable embedding of $\mathcal{M}$ gives us guarantees on data processing or inference algorithms in the measurement space [43]. Notably, it is shown in [73] that manifold learning (e.g., [18, 115, 129]) and dimensionality estimation algorithms (e.g., [68]) can be performed in the compressed space with nearly the same accuracy as in the original space. If the signal manifold were living in a high-dimensional space, then performing these algorithms in the compressed space (with dimension on the order of the dimension of the manifold) may result in massive speed increases.

The stable embedding result of Theorem 2.5.1 ensures that Euclidean distances between points on the manifold maintained in the measurement space. Not surprisingly, this also means that geodesic distances on the manifold are preserved:

Corollary 2.5.1. *[15] Let $\mathcal{M}$ be a manifold as in Theorem 2.5.1. If $\sqrt{\frac{N}{M}}\Phi$ is a stable embedding of $\mathcal{M}$ with conditioning δ, then for all $x, y \in \mathcal{M}$, we have*

$$(1-\delta)d_{\mathcal{M}}(x,y) \leq \sqrt{\frac{N}{M}} d_{\Phi(\mathcal{M})}(\Phi x, \Phi y) \leq (1+\delta)d_{\mathcal{M}}(x,y),$$

where $d_{\Phi(\mathcal{M})}(\Phi x, \Phi y)$ is the geodesic distance between the projected points on the image manifold $\Phi\mathcal{M}$.

Because geodesic distance are calculated as the integral sum of infinitesimal path length along the geodesic path, the preservation of Euclidean distance between points on the submanifold implies preservation of geodesic distances and thus, Corollary 2.5.1 follows (see [15] for more details).

It is useful to note that recent work on manifold learning [18, 41, 115, 129] also provides algorithms to project submanifolds onto lower dimensional Euclidean spaces that ensure some form of distance preservation. A difference between manifold learning and Theorem 2.5.1 is that these learning algorithms require (a large number of) training points on the manifold before the dimensionality reduction step can be done effectively.

CHAPTER III

BLOCK DIAGONAL MATRICES IN COMPRESSED SENSING

As discussed in the introduction, recent technological advances have enabled the sensing and storage of massive volumes of data from a dizzying array of sources. While access to such data has revolutionized fields such as signal processing, the limits of some computing and storage resources are being tested, and front-end signal acquisition devices are not always able to support the desire to measure in increasingly finer detail. To confront these challenges, many signal processing researchers in the CS community have begun investigating compressive linear operators $\Phi \in \mathbb{R}^{M \times N}$ for high resolution signals $x \in \mathbb{R}^N$ ($M < N$), either as a method for simple dimensionality reduction or as a model for novel data acquisition devices [25, 53]. Because of their universality and amenability to analysis, randomized compressive linear operators (i.e., random matrices with $M < N$) have drawn particular interest.

However, previous analysis of compressive operators to date has focused on dense matrices that require each measurement to be a weighted linear combination of all entries of x. Dense random matrices are often either impractical because of the resources required to store and work with a large unstructured matrix (e.g., one with i.i.d. entries), or unrealistic as models of acquisition devices with architectural constraints preventing such global data aggregation. For example, in a distributed sensing system, communication constraints may limit the dependence of each measurement to only a subset of the data. For a second example, applications involving streaming signals [9, 21] often have datarates that necessitate operating on local signal blocks rather than the entire signal simultaneously.

In such scenarios, the data may be divided naturally into discrete subsections (or blocks), with each block acquired via a local measurement operator. To see the implications of this, let us model a signal $x \in \mathbb{R}^{\widetilde{N}}$ as being partitioned into J blocks $x_1, x_2, \ldots, x_J \in \mathbb{R}^N$, and for each $j \in [J]$, suppose that a local measurement operator $\Phi_j : \mathbb{R}^N \to \mathbb{R}^{M_j}$ collects the measurements $y_j = \Phi_j x_j$. Concatenating all of the measurements into a vector $y \in \mathbb{R}^{\widetilde{M}}$, we then have

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_J \end{bmatrix}}_{y:\ \widetilde{M}\times 1} = \underbrace{\begin{bmatrix} \Phi_1 & & & \\ & \Phi_2 & & \\ & & \ddots & \\ & & & \Phi_J \end{bmatrix}}_{\Phi:\ \widetilde{M}\times\widetilde{N}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_J \end{bmatrix}}_{x:\ \widetilde{N}\times 1}, \tag{12}$$

where $\widetilde{M} = \sum_{j=1}^{J} M_j$ and $\widetilde{N} = NJ$. In cases such as these, we see that the overall measurement operator Φ will have a characteristic block diagonal structure. In some scenarios, the local measurement operator Φ_j may be unique for each block, and we say that the resulting Φ has a *Distinct Block Diagonal* (DBD) structure. In other scenarios, it may be appropriate or necessary to repeat a single operator across all blocks (such that $\Phi_1 = \Phi_2 = \cdots = \Phi_J$); we call the resulting Φ a *Repeated Block Diagonal* (RBD) matrix.

The focus of this chapter is a two-pronged analysis of such block diagonal operators. First in Section 3.1, we derive concentration of measure bounds for DBD and RBD matrices and explore the implications and utility of these bounds for the signal processing community.[1] Next in Section 3.2, we show that both DBD and RBD matrices can satisfy the RIP with high probability whenever the total number of measurements scales logarithmically with the ambient dimension and linearly with

[1]This work was performed in collaboration with Jae Young Park, Christopher J. Rozell, and Michael B. Wakin. JYP and HLY contributed equally to this work. Specifically, JYP was largely responsible for the general COM bounds and HLY made substantial contributions to the applications section of the work and derived the necessary bounds for frequency sparse signals. The results of this work are published in [106, 107, 117].

the sparsity and a coherence measure of the sparsity basis (that is distinct for each block diagonal matrix type).[2]

3.1 The Concentration of Measure Analysis

The theoretical analysis of random matrices often relies on the general notions that these matrices are well-behaved most of the time and that we can bound the probability with which they perform poorly. Frequently, these notions are formalized using some form of the concentration of measure phenomenon [87], a powerful characterization of the tendency of certain functions of high-dimensional random processes to concentrate sharply around their mean. As one important example of this phenomenon, it is known that for any fixed signal $x \in \mathbb{R}^N$, if Φ is an $M \times N$ matrix populated with i.i.d. random entries drawn from a suitable distribution, then with high probability Φ will approximately preserve the norm of x. More precisely, for many random distributions for Φ, the probability that $|\|\Phi x\|_2^2 - \|x\|_2^2|$ will exceed a small fraction of $\|x\|_2^2$ decays exponentially in the number of measurements M.

As we discussed in Sections 2.3 and 2.4 of this thesis, such concentration results have a number of favorable implications. Among these is the JL lemma (see Lemma 2.3.1), which states that when applied to a finite set of points $\mathcal{M} \subset \mathbb{R}^N$, a randomized compressive operator Φ can provide a stable, distance preserving embedding of $\mathcal{M}$ in the measurement space $\mathbb{R}^M$. This enables the efficient solution of a broad variety of signal processing problems by permitting these problems to be solved in the low-dimensional observation space (such as finding the nearest neighbor to a point x in a database $\mathcal{M}$). Such concentration results have also been used to prove that certain families of random matrices can satisfy the RIP (see Theorem 2.4.2).

Starting from the block diagonal matrix structure in (12), we derived concentration

[2]This work was performed in collaboration with Michael B. Wakin, Armin Eftekhari, and Christopher J. Rozell. AE and HLY contributed equally to this work, with both authors making substantial contributions to the development of the main theoretical results. The full results of this work are in a submitted manuscript [58], while initial results were presented in [143].

of measure bounds for DBD and RBD matrices. Specifically, we present concentration of measure bounds for DBD matrices populated with i.i.d. subgaussian random variables in Section 3.1.1 and for RBD matrices populated with i.i.d. Gaussian random variables in Section 3.1.2. In contrast to the signal agnostic concentration of measure bounds for i.i.d. dense matrices, these bounds are signal dependent; in particular, the probability of concentration depends on the "diversity" of the component signals $x_1, x_2, \ldots, x_J$ being well-matched to the measurement matrix (we make this precise in Sections 3.1.1 and 3.1.2). As our analytic discussion and supporting simulations show, these measures of diversity have clear intuitive interpretations and indicate that, for signals with the most favorable characteristics, the concentration of measure probability for block diagonal matrices can scale exactly as for an i.i.d. dense random matrix.

Sections 3.1.3 and 3.1.4 are devoted to a detailed investigation of the utility of these non-uniform concentration results for signal processing practitioners. Specifically, in Section 3.1.3, we extend our concentration results to formulate a modified version of the JL lemma appropriate for block diagonal matrices. We also explain how this lemma can be used to guarantee the performance of various compressive-domain signal inference and processing algorithms such as signal detection and estimation. Given the applicability of these results for providing performance guarantees in these tasks, a natural question is whether there are large classes of signals that have the diversity required to make block diagonal matrices perform well. In Section 3.1.4, we provide several examples of signal families that are particularly favorable for measurement via DBD or RBD matrices.

3.1.1 Distinct Block Diagonal Matrices

In this section, we state our concentration of measure results for DBD matrices and use simulations to demonstrate that our results do indeed capture the salient signal

characteristics that affect the concentration probability.

Before stating our result, we define the requisite notation. For a given signal $x \in \mathbb{R}^{\widetilde{N}}$ partitioned into J blocks of length N as in (12), we define a vector describing the energy distribution across the blocks of x:

$$\gamma = \gamma(x) := \begin{bmatrix} \|x_1\|_2^2 & \|x_2\|_2^2 & \cdots & \|x_J\|_2^2 \end{bmatrix}^T \in \mathbb{R}^J.$$

Also, letting $M_1, M_2, \ldots, M_J$ denote the number of measurements to be taken of each block, we define a $J \times J$ diagonal matrix containing these numbers along the diagonal:

$$\mathbf{M} := \operatorname{diag}(M_1, M_2, \ldots, M_J).$$

Finally, for a given signal $x \in \mathbb{R}^{\widetilde{N}}$ and measurement allocation $\mathbf{M}$, we define the quantities

$$\begin{aligned} \Gamma_2(x, \mathbf{M}) := \frac{\|\gamma\|_1^2}{\|\mathbf{M}^{-1/2}\gamma\|_2^2} &= \frac{\left(\sum_{j=1}^J \|x_j\|_2^2\right)^2}{\sum_{j=1}^J \frac{\|x_j\|_2^4}{M_j}}, \text{ and} \\ \Gamma_\infty(x, \mathbf{M}) := \frac{\|\gamma\|_1}{\|\mathbf{M}^{-1}\gamma\|_\infty} &= \frac{\sum_{j=1}^J \|x_j\|_2^2}{\max_j \frac{\|x_j\|_2^2}{M_j}}. \end{aligned} \tag{13}$$

Using this notation, our first result concerning the concentration of DBD matrices is captured in the following theorem.

Theorem 3.1.1. *Suppose $x \in \mathbb{R}^{\widetilde{N}}$, and for each $j \in [J]$ suppose that $M_j > 0$. Let ϕ denote a subgaussian random variable with mean 0, variance 1, and subgaussian norm $\|\phi\|_{\psi_2}$. Let $\{\Phi_j\}_{j=1}^J$ be random matrices drawn independently, where each Φ_j has size $M_j \times N$ and is populated with i.i.d. realizations of the renormalized random variable $\frac{\phi}{\sqrt{M_j}}$, and let Φ be a $\widetilde{M} \times \widetilde{N}$ DBD matrix composed of $\{\Phi_j\}_{j=1}^J$ as in (12). Then*

$$\begin{aligned} & \mathbb{P}\left\{ \left| \|\Phi x\|_2^2 - \|x\|_2^2 \right| > \epsilon \|x\|_2^2 \right\} \\ \leq\ & 2\exp\left(-C_1 \min\left\{ \frac{C_2^2 \epsilon^2}{\|\phi\|_{\psi_2}^4} \Gamma_2(x, \mathbf{M}), \frac{C_2 \epsilon}{\|\phi\|_{\psi_2}^2} \Gamma_\infty(x, \mathbf{M}) \right\} \right), \end{aligned} \tag{14}$$

where C_1 and C_2 are absolute constants.

Proof. See Appendix C.1. □

From the tail bound (14), it is easy to deduce that the concentration probability of interest decays exponentially as a function of $\epsilon^2\Gamma_2(x,\mathbf{M})$ in the case where $0 \leq \epsilon \leq \frac{\|\phi\|_{\psi_2}^2\Gamma_\infty(x,\mathbf{M})}{C_2\Gamma_2(x,\mathbf{M})}$ and exponentially as a function of $\epsilon\Gamma_\infty(x,\mathbf{M})$ in the case where $\epsilon > \frac{\|\phi\|_{\psi_2}^2\Gamma_\infty(x,\mathbf{M})}{C_2\Gamma_2(x,\mathbf{M})}$. One striking thing about Theorem 3.1.1 is that, in contrast to analogous concentration of measure results for dense matrices with i.i.d. subgaussian entries (4), the concentration rate depends explicitly on the signal x being measured.

To elaborate on this point, since we are frequently concerned in practice with applications where ϵ is small, let us focus on the first case of (14), when the concentration exponent scales with $\Gamma_2(x,\mathbf{M})$. In this case, we see that larger values of $\Gamma_2(x,\mathbf{M})$ promote sharper concentration of $\|\Phi x\|_2^2$ about its mean $\|x\|_2^2$. Using elementary inequalities relating the ℓ_1 and ℓ_2 norms, one can bound the range of possible Γ_2 values by

$$\min_j M_j \leq \Gamma_2(x,\mathbf{M}) \leq \widetilde{M} = \sum_{j=1}^{J} M_j.$$

The worst case, $\Gamma_2(x,\mathbf{M}) = \min_j M_j$, is achieved when all of the signal energy is concentrated into exactly one signal block where the fewest measurements are collected, i.e., when $\|x_j\|_2^2 = 0$ except for a single index $j' \in \{\arg\min_j M_j\}$ (where $\{\arg\min_j M_j\}$ is the set of indices where $\{M_j\}$ is minimum). In this case the DBD matrix exhibits significantly worse performance than a dense i.i.d. matrix of the same size $\widetilde{M} \times \widetilde{N}$, for which the concentration exponent would scale with $\widetilde{M}$. This makes intuitive sense, as this extreme case would correspond to only one block of the DBD matrix sensing all of the signal energy. On the other hand, the best case, $\Gamma_2(x,\mathbf{M}) = \widetilde{M}$, is achieved when the number of measurements collected for each block is proportional to the signal energy in that block. In other words, letting $\mathrm{diag}(\mathbf{M})$ represent the diagonal of $\mathbf{M}$, when $\mathrm{diag}(\mathbf{M}) \propto \gamma$ (i.e., when $\mathrm{diag}(\mathbf{M}) = C\gamma$ for some constant $C > 0$) the concentration exponent scales with $\widetilde{M}$ just as it would for a dense i.i.d. matrix of the

same size. This is in spite of the fact that the DBD matrix has many fewer nonzero elements.

The probability of concentration behaves similarly in the second case of (14), where the concentration exponent scales with $\Gamma_\infty(x, \mathbf{M})$. One can bound the range of possible Γ_∞ values by

$$\min_j M_j \leq \Gamma_\infty(x, \mathbf{M}) \leq \widetilde{M} = \sum_{j=1}^{J} M_j.$$

The lower bound is again achieved when $\|x_j\|_2^2 = 0$ except for a single index $j' \in \{\arg\min_j M_j\}$, and the upper bound again is achieved when $\mathrm{diag}(\mathbf{M}) \propto \gamma$.

The above discussion makes clear that the concentration performance of a DBD matrix can vary widely depending on the signal being measured. In particular, DBD matrices can perform as well as dense i.i.d. matrices if their measurement allocation is well matched to the energy distribution of the signal. Such a favorable event can occur either (i) by design, if a system designer has some operational knowledge of the energy distributions to expect, or (ii) by good fortune, if favorable signals happen to arrive that are well matched to a fixed system design. We note that even in the former situation when the general energy distribution across blocks is known, this does not imply that the designer has a priori knowledge of the signal being sensed. Furthermore, even when significant information about the signal (or a finite class of signals) is known, there may still be much to learn by actually measuring the signal. For example, Section 3.1.3 outlines several interesting signal inference problems that benefit from a norm-preservation guarantee for a known signal (or finite signal family). Also, in the second of these situations, it may not be unreasonable to expect that a fixed measurement allocation will be well matched to an unknown signal most of the time. For example, in Section 3.1.4 we describe several realistic signal classes that are favorably matched to fixed systems that have equal measurement allocations $(M_1 = M_2 = \cdots = M_J)$.

Two final comments are in order. First, while Theorem 3.1.1 was derived by considering all signal blocks to be of equal length N, one can see by a close examination of the proof that the same theorem in fact holds for signals partitioned into blocks of unequal lengths. Second, it is instructive to characterize the range of ϵ for which the two cases of Theorem 3.1.1 are relevant; we do so in the following lemma, which can be proved using standard manipulations of the ℓ_1, ℓ_2, and ℓ_∞ norms.

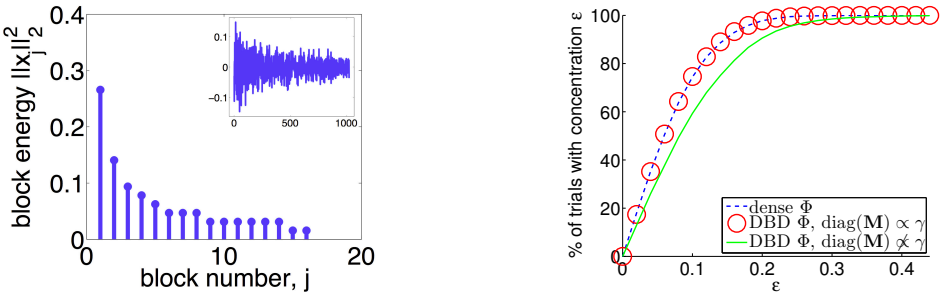
Lemma 3.1.1. *If $J \geq 2$, the point of demarcation between the two cases of Theorem 3.1.1 obeys*

$$\frac{\|\phi\|_{\psi_2}^2 \cdot 2(\sqrt{J}-1)}{C_2(J-1)} \frac{\min_j \sqrt{M_j}}{\max_j \sqrt{M_j}} \leq \frac{\|\phi\|_{\psi_2}^2 \, \Gamma_\infty(x, \mathbf{M})}{C_2 \Gamma_2(x, \mathbf{M})} \leq \frac{\|\phi\|_{\psi_2}^2}{C_2}.$$

Examining the bound above, we note that for $J \geq 2$ it holds that $\frac{2(\sqrt{J}-1)}{J-1} \geq \frac{1}{\sqrt{J}}$. Thus, as an example, when $M_1 = M_2 = \cdots = M_J$, the first ("small ϵ") case of Theorem 3.1.1 is guaranteed to at least permit $\epsilon \in \left[0, \frac{\|\phi\|_{\psi_2}^2}{C_2\sqrt{J}}\right]$. We note further that when the measurement matrix is well-matched to the signal characteristics, the first case of Theorem 3.1.1 permits ϵ as large as $\frac{\|\phi\|_{\psi_2}^2}{C_2}$, which is independent of J.

While the quantity $\Gamma_2(x, \mathbf{M})$ plays a critical role in our analytical tail bound (14), it is reasonable to ask whether this quantity actually plays a central role in the empirical concentration performance of DBD matrices. We explore this question with a series of simulations. To begin, we randomly construct a signal of length 1024 partitioned into $J = 16$ blocks of length $N = 64$. The energy distribution γ of the signal x is plotted in Figure 3(a) (and the signal x itself is plotted in the top right corner). For this simulation, to ensure $\text{diag}(\mathbf{M}) \propto \gamma$ with integer values for the M_j, we begin by constructing $\mathbf{M}$ (populated with integers) and then normalize each block of a randomly generated signal to set γ accordingly.

Fixing this signal x, we generate a series of 10000 random 64×1024 matrices Φ using zero-mean Gaussian random variables for the entries. In one case, the matrices



(a) (b)

Figure 3: (a) Test signal for concentration in a DBD matrix. The main panel plots the energy distribution γ when the signal is partitioned into $J = 16$ blocks of length $N = 64$; the subpanel plots the length-1024 signal x itself. (b) This figure shows the percentage of trials for which $(1-\epsilon) \leq \|\Phi x\|_2/\|x\|_2 \leq (1+\epsilon)$ as a function of ϵ for dense matrices, and for DBD matrices that either matched or not matched to the signal energy distribution.

are fully dense and the entries of each matrix have variance $1/64$. In another case, the matrices are DBD with $\mathrm{diag}(\mathbf{M}) \propto \gamma$ and the entries in each block have variance $1/M_j$. Thus, we have $\Gamma_2(x, \mathbf{M}) = \widetilde{M}$ and our Theorem 3.1.1 gives the same concentration bound for this DBD matrix as for the dense i.i.d. matrix of the same size. For each type of matrix, Figure 3(b) shows the percentage of trials for which $(1-\epsilon) \leq \|\Phi x\|_2/\|x\|_2 \leq (1+\epsilon)$ as a function of ϵ, and indeed, the curves for the dense and DBD matrices are indistinguishable.

Finally, we consider a third scenario in which we construct 10000 random 64×1024 DBD matrices as above but with an equal number of measurements in each block. In other words, we set all $M_j = 4$, and obtain measurement matrices that are no longer matched to the signal energy distribution. We quantify this mismatch by noting that $\Gamma_2(x, 4 \cdot I_{J\times J}) = 32.77 < \widetilde{M}$. Again, Figure 3(b) shows the concentration success probability over these 10000 random matrices. It is evident that these mismatched DBD matrices provide decidedly less sharp concentration of $\|\Phi x\|_2$.

3.1.2 Repeated Block Diagonal Matrices

In this section, we state our concentration of measure results for RBD matrices and again use simulations to demonstrate that our results do indeed capture the salient signal characteristics that affect the concentration probability. We also discuss connections between the concentration probabilities for the two matrix types.

We now turn our attention to the concentration performance of the more restricted RBD matrices. Before stating our result, let us again define the requisite notation. Given a signal $x \in \mathbb{R}^{\widetilde{N}}$ partitioned into J blocks of length N, we define the $J \times N$ matrix of concatenated signal blocks

$$X := [x_1 \mid x_2 \mid \cdots \mid x_J]^T, \tag{15}$$

and we denote the non-negative eigenvalues of the $N \times N$ symmetric matrix $A = X^T X$ as $\{\lambda_i\}_{i=1}^N$. We let $\lambda = \lambda(x) := [\lambda_1, \ldots, \lambda_N]^T \in \mathbb{R}^N$ be the vector composed of these eigenvalues. We let $M := M_1 = M_2 = \cdots = M_J$ denote the number of measurements to be taken in each block and the total number of measurements is now $\widetilde{M} = MJ$. Finally, for a given signal $x \in \mathbb{R}^{\widetilde{N}}$ and per-block measurement rate M, we define the quantities

$$\Lambda_2(x, M) := \frac{M\|\lambda\|_1^2}{\|\lambda\|_2^2} \quad \text{and} \quad \Lambda_\infty(x, M) := \frac{M\|\lambda\|_1}{\|\lambda\|_\infty}. \tag{16}$$

Equipped with this notation, our main result concerning the concentration of RBD matrices is as follows.

Theorem 3.1.2. *Suppose $x \in \mathbb{R}^{\widetilde{N}}$. Let $\widetilde{\Phi}$ be a random $M \times N$ matrix populated with i.i.d. zero-mean Gaussian entries having variance $\sigma^2 = \frac{1}{M}$, and let Φ be an $\widetilde{M} \times \widetilde{N}$ block diagonal matrix as defined in* (12)*, with $\Phi_j = \widetilde{\Phi}$ for all j. Then*

$$\begin{aligned} & \mathbb{P}\left\{ \left| \|\Phi x\|_2^2 - \|x\|_2^2 \right| > \epsilon \|x\|_2^2 \right\} \\ \leq \; & 2\exp\left(-C_1 \min\left\{ C_3^2 \epsilon^2 \Lambda_2(x, M),\; C_3 \epsilon \Lambda_\infty(x, M) \right\} \right), \end{aligned} \tag{17}$$

where C_1 and C_3 are absolute constants.

Proof. See Appendix C.2. □

From (17), one can deduce that the concentration probability of interest decays exponentially as a function of $\epsilon^2\Lambda_2(x,M)$ in the case where $0 \leq \epsilon \leq \frac{\Lambda_\infty(x,M)}{C_3\Lambda_2(x,M)}$ and exponentially as a function of $\epsilon\Lambda_\infty(x,M)$ in the case where $\epsilon > \frac{\Lambda_\infty(x,M)}{C_3\Lambda_2(x,M)}$. Thus, we see that the concentration rate again depends explicitly on the signal x being measured.

Again, since we are frequently concerned in practice with applications where ϵ is small, let us focus on the first case of (17), when the concentration exponent scales with $\Lambda_2(x,M)$. It follows from the standard relation between ℓ_1 and ℓ_2 norms that $M \leq \Lambda_2(x,M) \leq M\min(J,N)$. One extreme, $\Lambda_2(x,M) = M$, is achieved when $A = \sum_j x_j x_j^T$ has only one nonzero eigenvalue, implying that the blocks x_j are the same modulo a scaling factor. In this case, the RBD matrix exhibits significantly worse performance than a dense i.i.d. matrix of the same size $\widetilde{M} \times \widetilde{N}$, for which the concentration exponent would scale with $\widetilde{M}$ rather than M. However, this diminished performance is to be expected since the same $\widetilde{\Phi}$ is used to measure each identical signal block.

The other extreme, $\Lambda_2(x,M) = M\min(J,N)$ is favorable as long as $J \leq N$, in which case the concentration exponent scales with $\widetilde{M}$ just as it would for a dense i.i.d. matrix of the same size. For this case to occur, A must have J nonzero eigenvalues and they must all be equal. By noting that the nonzero eigenvalues of $A = X^TX$ are the same as those of the Grammian matrix $G = XX^T$, we conclude that this most favorable case can occur only when the signal blocks are mutually orthogonal and have the same energy. Alternatively, if the signal blocks span a K-dimensional subspace of $\mathbb{R}^N$ we will have $M \leq \Lambda_2(x,M) \leq MK$. All of this can also be seen by observing that calculating the eigenvalues of $A = X^TX$ is equivalent to running Principal Component Analysis (PCA) [64] on the matrix X comprised of the J signal blocks. Said another way, an RBD matrix performs as well as a dense i.i.d. matrix of

the same size when the signal has uniform energy distribution across its blocks (as in the DBD case) *and* has sufficient variation in the directions exhibited by the blocks.

We note that there is a close connection between the diversity measures $\Gamma_2(x, \mathbf{M})$ and $\Lambda_2(x, M)$ that is not apparent at first glance. For a fair comparison, we assume in this discussion that $\mathbf{M} := \text{diag}(M, M, \ldots, M)$. Now, note that $\|\lambda\|_1^2 = \|\gamma\|_1^2$ and also that

$$\|\lambda\|_2^2 = \|A\|_F^2 = \|XX^T\|_F^2 = \sum_{i=1}^{J} \|x_i\|_2^4 + 2\sum_{i>j}(x_i^T x_j)^2 = \|\gamma\|_2^2 + 2\sum_{i>j}(x_i^T x_j)^2.$$

Using these two relationships, we can rewrite $\Lambda_2(x, M)$ as

$$\Lambda_2(x, M) = \frac{M\|\lambda\|_1^2}{\|\lambda\|_2^2} = \frac{M\|\gamma\|_1^2}{\|\gamma\|_2^2 + 2\sum_{i>j}(x_i^T x_j)^2} \leq \frac{M\|\gamma\|_1^2}{\|\gamma\|_2^2} = \Gamma_2(x, \mathbf{M}). \tag{18}$$

From this relationship we see that Λ_2 and Γ_2 differ only by the additional inner-product term in the denominator of Λ_2, and we also see that $\Lambda_2 = \Gamma_2$ if and only if the signal blocks are mutually orthogonal. This more stringent condition for RBD matrices—requiring more intrinsic signal diversity—is expected given the more restricted structure of the RBD matrices.

While the quantity $\Lambda_2(x, M)$ plays a critical role in our analytical upper bound (17) on the concentration tail probabilities, it is reasonable to ask whether this quantity actually plays a central role in the empirical concentration performance of RBD matrices. We explore this question with a series of simulations. To begin, we randomly construct a signal of length 1024 partitioned into $J = 16$ blocks of length $N = 64$, and we perform Gram-Schmidt orthogonalization to ensure that the J blocks are mutually orthogonal and have equal energy. The nonzero eigenvalues of $A = X^T X$ are shown in the plot of λ in Figure 4(a) (and the signal x itself, denoted "Sig. 1", is plotted in the top left corner).

As we have discussed above, for signals such as Sig. 1 we should have $\Lambda_2(x, M) =$

$\widetilde{M}$, and Theorem 3.1.2 suggests that an RBD matrix can achieve the same concentration rate as a dense i.i.d. matrix of the same size. Fixing this signal, we generate a series of 10000 random 64×1024 matrices Φ populated with zero-mean Gaussian random variables. In one case, the matrices are dense and each entry has variance $1/64$. In another case, the matrices are RBD, with a single 4×64 block repeated along the main diagonal, comprised of i.i.d. Gaussian entries with variance $\frac{1}{4}$. For each type of matrix, Figure 4(c) shows the percentage of trials for which $(1-\epsilon) \leq \|\Phi x\|_2/\|x\|_2 \leq (1+\epsilon)$ as a function of ϵ. As anticipated, we can see that the curves for the dense and RBD matrices are indistinguishable.

In contrast, we also construct a second signal x (denoted "Sig. 2") that has equal energy between the blocks but has non-orthogonal components (resulting in non-uniform λ); see Figure 4(b). This signal was constructed to ensure that the sorted entries of λ exhibit an exponential decay. Due to the non-orthogonality of the signal blocks, we see for this signal that $\Lambda_2(x, M) = 21.3284$ which is approximately 3 times less than the best possible value of $\widetilde{M} = 64$. Consequently, Theorem 3.1.2 provides a much weaker concentration exponent when this signal is measured using an RBD matrix than when it is measured using a dense i.i.d. matrix. As shown in Figure 4(c), we see that the concentration performance of the full dense matrix is agnostic to this new signal structure, while the concentration is clearly not as sharp for the RBD matrix.

3.1.3 Applications

A concentration of measure inequality, despite nominally pertaining to the norm preservation of a single signal, can lead to a number of guarantees for problems involving multi-signal embeddings and signal discrimination. In this section, we extend our concentration bounds to formulate a modified version of the JL lemma appropriate for block diagonal matrices. We also survey a collection of compressive-domain

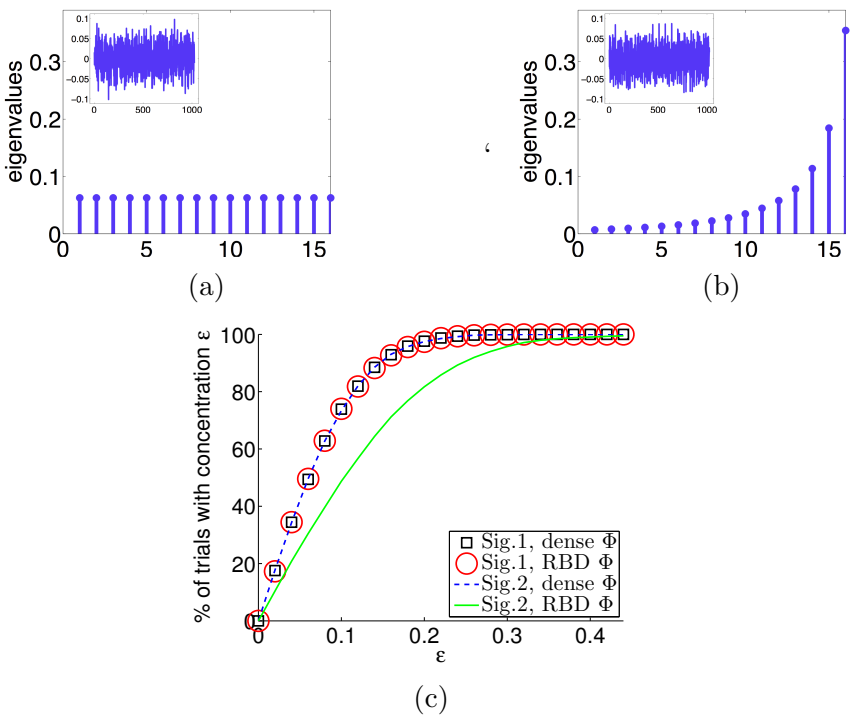


Figure 4: **(a),(b) Test signals for concentration in a RBD matrix. The main panels plot the eigenvalue distributions λ for Sig. 1 and Sig. 2, respectively, when partitioned into $J = 16$ blocks of length $N = 64$; the subpanels plot the length-1024 signals themselves. (c) This figure shows the percentage of trials for which $(1 - \epsilon) \leq \|\Phi x\|_2 / \|x\|_2 \leq (1 + \epsilon)$ as a function of ϵ for dense matrices and for RBD matrices.**

inference problems (such as detection and classification) in which such a result can be leveraged. For simplicity we will focus on DBD matrices in this section, but parallel results can be derived in each case for RBD matrices. Given the nonuniform nature of our concentration bounds, the performance of algorithms for solving these problems will depend on the signals under consideration, and so, in Section 3.1.4 we provide several examples of signal classes that are particularly favorable for measurement via DBD or RBD matrices.

3.1.3.1 Stable Embeddings and the Johnson-Lindenstrauss Lemma

For a given signal $x \in \mathbb{R}^{\widetilde{N}}$ and measurement allocation $\mathbf{M}$, let us define $\widetilde{\Gamma}_2(x, \mathbf{M}) := \frac{\Gamma_2(x,\mathbf{M})}{\widetilde{M}}$ and $\widetilde{\Gamma}_\infty(x, \mathbf{M}) := \frac{\Gamma_\infty(x,\mathbf{M})}{\widetilde{M}}$. Note that both quantities are upper bounded by

1, with equality achieved for signals best matched to $\mathbf{M}$ as discussed in Section 3.1.1. Here, we overload the definition of a stable embedding by saying that Φ is a stable embedding of (U, V) with conditioning δ if for all $u \in U$ and $v \in V$,

$$(1-\delta)\|u-v\|_2^2 \leq \|\Phi(u-v)\|_2^2 \leq (1+\delta)\|u-v\|_2^2.$$

Using this notation, Theorem 3.1.1 allows us to immediately formulate a version of the JL lemma appropriate for DBD matrices.

Theorem 3.1.3. *Let U, V be two finite subsets of $\mathbb{R}^{\widetilde{N}}$, let Φ be a randomly generated DBD matrix as described in Theorem 3.1.1 with measurement allocation $\mathbf{M}$, and let $\rho \in (0,1)$ be fixed. If*

$$\widetilde{M} \geq \frac{\log|U| + \log|V| + \log(2/\rho)}{C_1 \min\left\{ \frac{C_2^2\delta^2}{\|\phi\|_{\psi_2}^4} \min_{u\in U, v\in V} \widetilde{\Gamma}_2(u-v, \mathbf{M}), \frac{C_2\delta}{\|\phi\|_{\psi_2}^2} \min_{u\in U, v\in V} \widetilde{\Gamma}_\infty(u-v, \mathbf{M}) \right\}}, \quad (19)$$

then with probability exceeding $1-\rho$, Φ will provide a stable embedding of (U, V) with conditioning δ. Alternatively, under the same conditions and with probability exceeding $1-\rho$, the matrix Φ will provide a stable embedding of (U, V) with conditioning

$$\begin{aligned} \widetilde{\delta}(U, V, \mathbf{M}, \rho) \;:=\; & \frac{\|\phi\|_{\psi_2}^2}{C_2} \max\left\{ \sqrt{\frac{\log|U| + \log|V| + \log(2/\rho)}{C_1 \min_{u\in U, v\in V} \Gamma_2(u-v, \mathbf{M})}}, \cdots \right. \\ & \left. \cdots \frac{\log|U| + \log|V| + \log(2/\rho)}{C_1 \min_{u\in U, v\in V} \Gamma_\infty(u-v, \mathbf{M})} \right\}. \end{aligned} \quad (20)$$

Proof. Taking the union bound over all $u \in U$ and $v \in V$ and using (14), we then know that the desired (near) isometry holds over all difference vectors $u - v$ except with probability bounded by

$$\begin{aligned} 2\,|U|\,|V| \exp\left(-C_1 \min\left\{ \frac{C_2^2\delta^2}{\|\phi\|_{\psi_2}^4} \min_{u\in U, v\in V} \Gamma_2(u-v, \mathbf{M}), \cdots \right. \right. \\ \left. \left. \cdots \frac{C_2\delta}{\|\phi\|_{\psi_2}^2} \min_{u\in U, v\in V} \Gamma_\infty(u-v, \mathbf{M}) \right\} \right). \end{aligned} \quad (21)$$

Ensuring that (19) holds guarantees that the above failure probability is less than ρ. The bound in (20) follows from (19) and the observation that $\min(a\delta^2, b\delta) = c$ implies that $\delta = \max\left(\sqrt{\frac{c}{a}}, \frac{c}{b}\right)$. □

Similar theorems can be formulated for RBD matrices, as well as for stable embeddings of a signal subspace rather than just a finite family of signals. Equation (19) gives a lower bound on the total number of measurements to guarantee a stable embedding with conditioning δ. One can see that the denominator on the right-hand side will scale with $\delta^2 \cdot \min_{u \in U, v \in V} \widetilde{\Gamma}_2(u-v, \mathbf{M})$ when $0 \leq \delta \leq \frac{\|\phi\|_{\psi_2}^2 \min_{u \in U, v \in V} \widetilde{\Gamma}_\infty(u-v, \mathbf{M})}{C_2 \min_{u \in U, v \in V} \widetilde{\Gamma}_2(u-v, \mathbf{M})}$ and with $\delta \cdot \min_{u \in U, v \in V} \widetilde{\Gamma}_\infty(u-v, \mathbf{M})$ when $\delta > \frac{\|\phi\|_{\psi_2}^2 \min_{u \in U, v \in V} \widetilde{\Gamma}_\infty(u-v, \mathbf{M})}{C_2 \min_{u \in U, v \in V} \widetilde{\Gamma}_2(u-v, \mathbf{M})}$. Thus, focusing just on cases where δ is small, in order to guarantee a stable embedding with a moderate number of measurements, we require $\widetilde{\Gamma}_2(u-v, \mathbf{M})$ to be sufficiently close to 1 over all $u \in U$ and $v \in V$. Equivalently, if $\Gamma_2(u-v, \mathbf{M})$ is uniformly close to $\widetilde{M}$ over all $u \in U$ and $v \in V$, the conditioning $\widetilde{\delta}$ provided in (20) is comparable to what would be achieved with a dense i.i.d. random matrix of the same size. In Section 3.1.4, we provide several examples of signal classes of U and V for which it may be reasonable to expect such uniformly favorable Γ_2 (or Λ_2) values.

The attentive reader may notice that the failure probability in (21) can be loose, since we have bounded the sum of the individual failure probabilities by $|U|\,|V|$ times the worse case failure probability. Due to the exponential form of these probabilities, however, it seems that the worse case probability—even if it is rare—will typically dominate this sum. Therefore, in most applications we do not expect that it is possible to significantly improve over the bounds provided in (21) and thus (19). Unfortunately, it appears that this fact would forbid the derivation of a sharp RIP bound for block diagonal matrices via the elementary covering arguments seen in the proof of Theorem 2.4.2. Nonetheless, we shall show in Section 3.2 that we are able to derive these RIP bounds by instead using recent techniques for bounding the suprema of chaos processes.

Indeed, ensuring a stable embedding of even a finite signal class is very useful for guaranteeing the performance of many types of compressive-domain signal inference and processing algorithms. In the following subsections, we explore two canonical

tasks (detection and classification) in detail to show how signal characteristics affect one's ability to solve these problems using measurements acquired via a block diagonal matrix. Performing tasks such as these directly in the measurement space not only reduces the data acquisition burden but can also reduce the computational burden far below what is required to solve these problems in the high dimensional ambient signal space.

We briefly note that there are several other tasks (aside from detection and classification) that can be performed in the measurement space when a block diagonal matrix provides a stable embedding of a finite signal family. For one example, when a block diagonal matrix Φ provides a stable embedding of $(\mathcal{S}, \{x\})$ for some signal database $\mathcal{S}$ and query signal x, it is possible to solve the approximate nearest neighbor problem [76] (finding the closest point in $\mathcal{S}$ to x) in the compressed domain without much loss of precision. Another potential application in compressive signal processing involves a simple compressive-domain linear estimator [43]. When Φ provides a stable embedding of $(\mathcal{L}, \mathcal{X} \cup -\mathcal{X})$ for some sets $\mathcal{L}$ and $\mathcal{X}$, then for any $\ell \in \mathcal{L}$ and $x \in \mathcal{X}$, we can estimate the value of $\langle \ell, x \rangle$ from the measurements $\langle \Phi\ell, \Phi x \rangle$. Signal families $\mathcal{L}$ and $\mathcal{X}$ whose sum and difference vectors $\ell \pm x$ have favorable Γ_2 values will have favorable and predictable estimation performance. Finally, a similar result also discussed in [43] shows that *filtering* vectors in order to separate signal and interference subspaces is possible when the difference vectors between these subspaces are stably embedded by Φ.

Before concluding this subsection, we note that while the concentration of measure analysis (of RBD matrices in particular) forbids us to achieve good RIP results via covering arguments, we have been slightly more successful in applying this method for the RIP of Toeplitz matrices. In a technical report [145], we obtained a measurement rate that has a quadratic scaling with sparsity by observing that compressive Toeplitz matrices are in fact related to RBD matrices. We will not discuss this result here

as the state-of-the-art RIP result for compressive Toeplitz matrices has achieved a linear scaling with sparsity [83]. However, we note that we shall leverage this duality between RBD matrices and compressive Toeplitz matrices again when we establish the RIP result for block diagonal matrices in Section 3.2.

3.1.3.2 Signal Detection in the Compressed Domain

While the canonical CS results center mostly around reconstructing signals from compressive measurements, there is a growing interest in forgoing this recovery process and answering certain signal processing questions directly in the compressed domain. One such problem that can be solved is binary detection, where one must decide whether or not a known template signal was present when the noisy compressive measurements were collected [43, 44, 72, 119]. In particular, let $s \in \mathbb{R}^{\tilde{N}}$ denote a known signal, and suppose that from the measurement vector y, we wish to decide between two hypotheses:

$$\mathcal{H}_0 : \ y = z \quad \text{or} \quad \mathcal{H}_1 : \ y = \Phi s + z,$$

where Φ is a known compressive measurement matrix, and z is a vector of i.i.d. zero-mean Gaussian noise with variance σ^2. If one were designing a measurement matrix specifically for the purpose of detecting this signal, then the optimal choice of Φ would be the matched filter, i.e., $\Phi = s^T$. However, implementing a measurement system that is designed specifically for a particular s restricts its capabilities to detecting that signal only, which could require a hardware modification every time s changes. A more generic approach would be to design Φ randomly (perhaps with a block diagonal structure out of necessity or due to efficiency considerations) and then use the acquired measurements y to test for one or more candidate signals s.

Given the measurements y, a Neyman-Pearson (NP) optimal detector [43] maximizes the probability of detection, $P_D := \mathbb{P}\{\mathcal{H}_1 \text{ chosen } | \mathcal{H}_1 \text{ is true}\}$, subject to a given limitation on the probability of a false alarm, $P_F = \mathbb{P}\{\mathcal{H}_1 \text{ chosen } | \mathcal{H}_0 \text{ is true}\}$.

The optimal decision for this problem is made based on whether or not the sufficient statistic $t := y^T \Phi s$ surpasses a threshold κ, i.e., $t \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \kappa$, where κ is chosen to meet the constraint $P_F \leq \alpha$ for a specified α. As can be seen from the analysis in [43], the performance of such a detector depends on $\|\Phi s\|_2$. In effect, if Φ "loses" signal energy for some signals the detector performance will suffer, and if Φ "amplifies" signal energy for some signals the detector performance will improve. A stable embedding of any signal the detector may encounter, however, guarantees consistent performance of the detector.

Theorem 3.1.4. *Suppose $\mathcal{S}$ is a finite set of signals and let Φ be a randomly generated DBD matrix as described in Theorem 3.1.1 with a number of measurements denoted by the matrix $\mathbf{M}$. Fix $0 < \rho < 1$ and pick α such that $P_F \leq \alpha$. Then with probability exceeding $1 - \rho$, any signal $s \in \mathcal{S}$ can be detected with probability of detection bounded by*

$$Q\left(Q^{-1}(\alpha) - \sqrt{1 - \widetilde{\delta}(\mathcal{S}, \{0\}, \mathbf{M}, \rho)}\sqrt{\frac{\|s\|_2^2}{\sigma^2}}\right) \leq P_D(\alpha) \leq \cdots$$
$$\cdots Q\left(Q^{-1}(\alpha) - \sqrt{1 + \widetilde{\delta}(\mathcal{S}, \{0\}, \mathbf{M}, \rho)}\sqrt{\frac{\|s\|_2^2}{\sigma^2}}\right),$$

where $Q(\alpha) = \frac{1}{\sqrt{2\pi}} \int_\alpha^\infty e^{-\frac{u^2}{2}} du$ and where $\widetilde{\delta}(\mathcal{S}, \{0\}, \mathbf{M}, \rho)$ is as defined in (20).

The proof of this theorem follows by combining the fact that

$$P_D(\alpha) = Q\left(Q^{-1}(\alpha) - \frac{\|\Phi s\|_2}{\sigma}\right)$$

(see [43]) with (20) and the monotonicity of the function $Q(\cdot)$. While achieving the best possible P_D for a given P_F is of course desirable for a detector, another very important consideration for a system designer is the reliability and consistency of that system. Large fluctuations in performance make it difficult to ascribe meaning to a particular detection result and to take action with a certain level of confidence. The theorem above tells us that the consistency of the detector performance is tied

to how reliably Φ preserves the norm of signals in $\mathcal{S}$. Examining this relationship, it is clear that more favorable values of $\Gamma_2(s, \mathbf{M})$ for a signal class result in tighter bounds on $P_D(\alpha)$ and therefore in stronger consistency guarantees for the detector.

To illustrate this fact with an example, we create a single DBD measurement matrix $\Phi \in \mathbb{R}^{\widetilde{M} \times \widetilde{N}}$ having an equal number of measurements $M_j = M$ per block. We take $M = 4$, $J = 16$ and $N = 64$, and we draw the nonzero entries of Φ as i.i.d. Gaussian random variables with variance $1/M$. We test the detection performance of this matrix by drawing 10000 unit-norm test signals randomly from two classes: $\mathcal{S}_1$, in which signals have uniform energy across their blocks, and $\mathcal{S}_2$, in which signals have decaying energy across their blocks. We choose the noise variance σ^2 such that each test signal s has a constant signal-to-noise ratio of $10 \log_{10}\left(\frac{\|s\|_2^2}{\sigma^2}\right) = 8\text{dB}$. Because of the construction of $\mathcal{S}_1$, $\Gamma_2(s, \mathbf{M})$ attains its maximum value of $\widetilde{M}$ for all signals $s \in \mathcal{S}_1$, resulting a small conditioning $\widetilde{\delta}$ and a tight bound on P_D. In contrast, $\mathcal{S}_2$ will have a smaller value of $\Gamma_2(s, \mathbf{M})$, resulting in larger values of $\widetilde{\delta}$ and much looser bounds on P_D. We choose the maximum probability of failure to be $\alpha = 0.1$ and use the equation $P_D(\alpha) = Q\left(Q^{-1}(\alpha) - \frac{\|\Phi s\|_2}{\sigma}\right)$ to calculate the expected P_D for each signal.

Figure 5 shows the histogram of P_D for the signals in $\mathcal{S}_1$ and $\mathcal{S}_2$ when measured with a DBD matrix. We see that for the uniform energy signals in $\mathcal{S}_1$, the detector performance is indeed tightly clustered around $P_D = 0.9$; one can see that this is the point of concentration predicted by Theorem 3.1.4 since $Q\left(Q^{-1}(0.1) - \sqrt{10^{8/10}}\right) \approx 0.8907$. Thus for this class of signals, the detector performance is consistent and we can be assured of a favorable P_D when using the detector for all signals in $\mathcal{S}_1$. However, when using the DBD matrix on the signal class $\mathcal{S}_2$, the P_D values are widely spread out compared to those for $\mathcal{S}_1$, despite the fact that the average P_D is nearly the same. Although some individual signals may have above average performance because the measurement matrix happened to amplify their energies, other signals may have

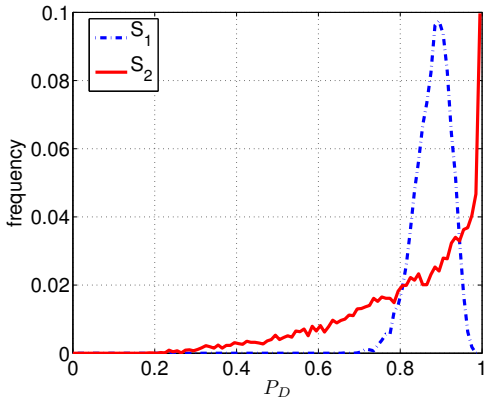


Figure 5: Histogram of P_D for 10000 signals with uniform energy across blocks (signal class $\mathcal{S}_1$) and for 10000 signals with decaying energy across blocks (signal class $\mathcal{S}_2$) when measured with a DBD matrix. The compressive NP detector has the constraint $P_F \leq \alpha = 0.1$.

very poor performance because the measurement matrix significantly attenuated their energies. Thus this experiment shows how the signal statistics affect the performance reliability in compressive detection tasks when the measurements matrices have block diagonal structure.

3.1.3.3 Classification in the Compressed Domain

Rather than determining the presence or absence of a fixed candidate signal, some scenarios may require the classification of a signal among multiple hypotheses [43, 44]. In particular, let $s_1, s_2, \ldots, s_R \in \mathbb{R}^{\tilde{N}}$ denote known signals, and suppose that from the measurement vector y, we wish to decide between R hypotheses:

$$\mathcal{H}_i : y = \Phi s_i + z, \text{ for } i = 1, 2, \ldots, R,$$

where Φ is a known compressive measurement matrix, and z is a vector of i.i.d. zero-mean Gaussian noise with variance σ^2.

It is straightforward to show that when each hypothesis is equally likely, the classifier with minimum probability of error selects the hypothesis that minimizes the sufficient statistic $t_i := \|y - \Phi s_i\|_2^2$. As can be imagined, the performance of such a classifier depends on how well Φ preserves pairwise distances among the signals $\{s_i\}$. If

a situation were to occur where $\frac{\|\Phi s_p - \Phi s_q\|_2}{\|s_p - s_q\|_2}$ was small for some p, q, then s_p could easily be mistaken for s_q in the measurements y. Therefore, having a stable embedding can again be particularly useful for guaranteeing consistent and predictable performance.

Theorem 3.1.5. *Let $\mathcal{S}$ denote a fixed set of signals with $|\mathcal{S}| = R < \infty$ and fix $0 < \rho < 1$. Suppose Φ is a randomly generated DBD matrix as described in Theorem 3.1.1 with a number of measurements denoted by the matrix $\mathbf{M}$. Assume we receive the measurements $y = \Phi s_{i^*} + z$ for some $i^* \in [R]$ and $z \sim \mathcal{N}(0, \sigma^2 I)$. Then, with probability at least*

$$1 - \left(\frac{R-1}{2}\right) \exp\left\{ -\frac{d^2\left(1 - \widetilde{\delta}(\mathcal{S}, \mathcal{S}, \mathbf{M}, \rho)\right)}{8\sigma^2} \right\} - 2\rho,$$

we have $i^* = \arg\min_{i \in [R]} t_i$ and thus the signal s_{i^*} can be correctly classified. Here $d := \min_{i,j} \|s_i - s_j\|_2$ denotes the minimum separation among the signals in $\mathcal{S}$ and $\widetilde{\delta}(\mathcal{S}, \mathcal{S}, \mathbf{M}, \rho)$ is as defined in (20).

The proof of this theorem again follows by combining bounds from [43] with (20). From this theorem it follows that, if Φ is a block diagonal matrix, signal families $\mathcal{S}$ whose difference vectors $s_p - s_q$ have favorable Γ_2 values will have consistent and predictable classification performance.

The following simulation demonstrates the potential for predictable classification of signals acquired using compressive block diagonal matrices. We again consider DBD matrices having an equal number of measurements $M_j = M$ per block, and we consider signals having $J = 16$ blocks of length $N = 64$. We first create a favorable class of unit-norm signals $\mathcal{S}_1$ with $R = J$ elements such that each signal has just 4 nonzero DFT coefficients at randomly chosen frequencies. To ensure that the signals are real, we restrict the coefficients on conjugate pairs of frequencies to have complex conjugate values. We also ensure that no frequencies are repeated amongst the signals in $\mathcal{S}_1$. As we show in Section 3.1.4.2, frequency sparse signals with

randomly selected frequency support will have large Γ_2 values with high probability; therefore the difference of any two signals from $\mathcal{S}_1$ will also have a large Γ_2 value with high probability. We also create a second class of unit-norm signals $\mathcal{S}_2$ with $R = J$ elements such that each signal s_r for $r = 1, 2, \ldots, R$ has nonzero (and randomly selected) values only on its r-th block and is zero everywhere else. Difference signals from this class will have small Γ_2 values since their energies across the blocks are not uniform.

For each M ranging from 1 to 20, we create 1000 instances of a random DBD matrix Φ of size $\widetilde{M} \times \widetilde{N}$. For each Φ and for each signal class $\mathcal{S}_1$ and $\mathcal{S}_2$, we identify the indices i_1, i_2 that minimize $\|\Phi s_{i_1} - \Phi s_{i_2}\|_2$. The signals corresponding to s_{i_1} and s_{i_2} will be among the most difficult to classify since they each have a close neighbor (either Φs_{i_2} or Φs_{i_1}, respectively) after projection by Φ. Then for each of these signals $\{s_{i_j}\}_{j=1,2}$, we create 1000 noisy measurement vectors $y = \Phi s_{i_j} + z$ with $z \sim \mathcal{N}(0, \sigma^2 I)$ and with σ chosen such that $10 \log_{10}\left(\frac{d^2}{\sigma^2}\right) = 15\text{dB}$. Finally, we let $p = \arg\min t_i$ be the output of the classifier and calculate the probability of misclassification, $P_E(M)$, for each M as the proportion of occurrences of $p \neq i_1$ or $p \neq i_2$, respectively, over the combined 1000 instances of noise z and 1000 instances of Φ.

Figure 6 plots $P_E(M)$ for both classes of signals. As expected, the curve for $\mathcal{S}_1$ is lower than that for $\mathcal{S}_2$ since the signals in $\mathcal{S}_1$ are expected to have a stable embedding with a tighter conditioning. Both curves also show a decreasing trend for increasing M (although it is much more obvious for signal class $\mathcal{S}_2$) as should be expected. Lastly, we see that $P_E(M)$ saturates at a certain level with increasing M. This is also predicted by Theorem 3.1.5, where the smallest upper bound that can be provided for P_E is given by $\frac{R-1}{2} \exp\left(-\frac{d^2}{8\sigma^2}\right) > 0$. With the parameters used in this experiment, it can be calculated that the smallest theoretical upper bound for P_E is approximately 0.144. This shows that Theorem 3.1.5 may be slightly pessimistic.

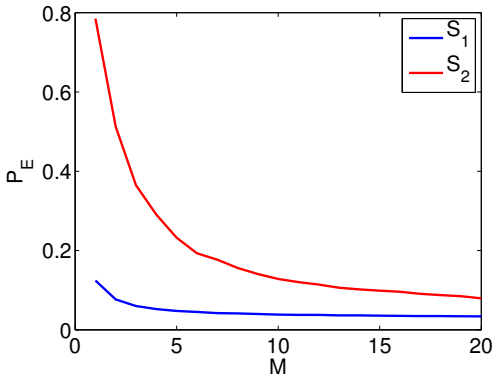


Figure 6: Plots of the probability of misclassification P_E over a range of values of $M = 1, \cdots, 20$. The first class of signals $\mathcal{S}_1$ are sparse in the frequency domain. The second class of signals $\mathcal{S}_2$ are nonzero only on a single block in the time domain. While P_E decreases with increasing M for both classes of signals, classification performs better for the signals in $\mathcal{S}_1$, which are more amenable to a stable embedding with a DBD matrix.

3.1.4 Favorable Signal Classes

The various compressive signal processing guarantees presented in Section 3.1.3 are built upon the premise that a DBD matrix provides a stable embedding of the signals of interest; as we have noted, these arguments can be extended to RBD matrices as well. Our analysis has also indicated that such stable embeddings are most easily realized with matrices that are well matched to the energy distribution (and sometimes orthogonality) of the signal blocks. In many applications, however—perhaps for the sake of generality, or because little advance knowledge of the signals is available—it may be most natural to use a fixed and equal allocation of measurements. Fortunately, there are a number of interesting signal families (and, in some cases, the corresponding difference signals) that provide favorable Γ_2 values for "uniform" DBD matrices where all M_j are equal (for all j, we suppose $M_j = M$ for some M) and in some cases also provide favorable Λ_2 values for RBD matrices. In this section we survey several such examples.

3.1.4.1 Delay Networks and Multiview Imaging

One favorable signal class for uniform DBD and RBD matrices can occasionally arise in certain sensor network or multi-view imaging scenarios where signals with steeply decaying autocorrelation functions are measured under small perturbations. Consider for example a distributed sensor network of J sensors where we would like to detect the presence of a known signal given the observations from each sensor. Due to limited resources, each sensor uses random measurement matrices $\Phi_1, \Phi_2, \ldots, \Phi_J$ to efficiently capture the underlying information with only a few random projections. Suppose that the received signals $x_1, x_2, \ldots, x_J \in \mathbb{R}^N$ represent observations of some common known underlying prototype signal $w \in \mathbb{R}^N$. However, due to the configurations of the sensors, these observations occur with different delays or translations. More formally, we might consider the one-dimensional delay parameters $\delta_1, \delta_2, \ldots, \delta_J \in \mathbb{Z}$ and have that for each j, $x_j(n) = w(n - \delta_j)$. Then, denoting the measurements at sensor j as $y_j = \Phi_j x_j$ it is straightforward to see that the overall system of equations can be represented with a DBD matrix, or when $\Phi_1 = \Phi_2 = \cdots = \Phi_J$ with an RBD matrix.

Assuming w is suitably zero-padded so that border and truncation artifacts can be neglected, we will have $\|x_j\|_2 = \|w\|_2$ for all x_j; this gives $\Gamma_2([x_1^T \; x_2^T \; \cdots x_J^T]^T, \mathbf{M}) = \widetilde{M}$, which is the ideal case for observation with a uniform DBD matrix. This suggests that the outputs from distributed network systems can be highly amenable to the sort of compressive-domain signal inference and processing tasks described Section 3.1.3.

Moreover, the correlations among the components x_j can be characterized in terms of the autocorrelation function R_w of w: we will have $\langle x_i, x_j \rangle = \sum_{n=1}^{N} x_i(n) x_j(n) = \sum_{n=1}^{N} w(n-\delta_i) w(n-\delta_j)$, which neglecting border and truncation artifacts will simply equal $R_w(|\delta_i - \delta_j|)$. Therefore, signals w that exhibit strong decay in their autocorrelation function will be natural candidates for observation with RBD matrices as well.

For example, equation (18) gives

$$\Lambda_2([x_1^T\ x_2^T\ \cdots x_J^T]^T, M) = \frac{MJ^2\|w\|_2^4}{J\|w\|_2^4 + 2\sum_{i>j} R_w(|\delta_i - \delta_j|)^2}.$$

When $R_w(|\delta_i - \delta_j|)$ is small for most i and j, the quantity $\Lambda_2([x_1^T\ x_2^T\ \cdots x_J^T]^T, M)$ is near its optimal value of $\widetilde{M}$.

3.1.4.2 Frequency Sparse Signals

Signals having sparse frequency spectra arise in many different applications involving communications intelligence systems and RF sensor networks. Based on time-frequency uncertainty principles and the well-known incoherence of sinusoids and the canonical basis (i.e., "spikes and sines") [28, 131], it is natural to expect that most signals that are sparse in the frequency domain should have their energy distributed relatively uniformly across blocks in the time domain. In the following theorem, we make formal the notion that frequency sparse signals are indeed likely to be favorable for measurement via uniform DBD matrices, producing values of $\Gamma_2(x, \mathbf{M})$ within a log factor of its maximum possible.

Theorem 3.1.6. *Let $N, \beta > 1$ be fixed, suppose $\widetilde{N} > 512$, and let $\mathbf{M} = \mathrm{diag}(M, M, \ldots, M)$ be a DBD measurement allocation with M fixed. Let $\Omega \subset [1,\ \widetilde{N}]$ be of size $S \leq N$ generated uniformly at random. Then with probability at least $1 - O(J(\log(\widetilde{N}))^{1/2}(\widetilde{N})^{-\beta})$,[3] every signal $x \in \mathbb{C}^{\widetilde{N}}$ whose DFT coefficients are supported on Ω will have:*[4]

$$\Gamma_2(x, \mathbf{M}) \geq \widetilde{M} \cdot \min\left\{ \frac{0.0779}{(\beta+1)\log(\widetilde{N})}, \frac{1}{\left(\sqrt{6(\beta+1)\log\widetilde{N}} + \frac{(\log\widetilde{N})^2}{N}\right)^2} \right\}. \qquad (22)$$

[3] *The $O(\cdot)$ notation is with respect to $\widetilde{N}$. With the component length N fixed, this means that only the number of blocks J is growing with increasing $\widetilde{N}$.*

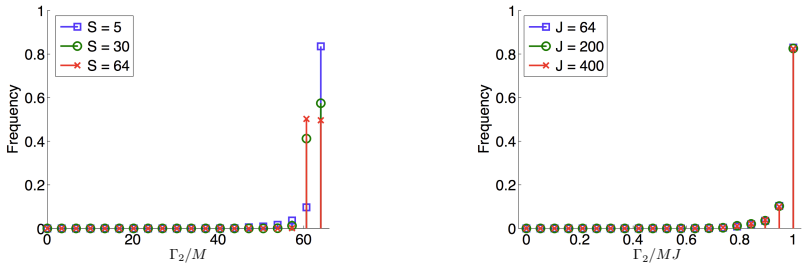
[4] *We consider complex-valued signals for simplicity and clarity in the exposition. A result with the same spirit that holds with high probability can be derived for strictly real-valued signals, but this comes at the cost of a more cumbersome derivation.*

Proof. See Appendix C.3. □

Note that as $\widetilde{N}$ grows, the lower bound on $\Gamma_2(x, \mathbf{M})$ scales as $\frac{\widetilde{M}}{\frac{1}{N^2}\log^4(\widetilde{N})}$, which (treating the fixed value N as a constant) is within $\log^4(\widetilde{N})$ of its maximum possible value of $\widetilde{M}$. Thus the concentration exponent for *most* frequency sparse signals when measured by a uniform DBD matrix will scale with $\epsilon^2\widetilde{M}/\log^4(\widetilde{N})$ for small ϵ. Also note that the failure probability in the above theorem can be bounded by $O(\frac{1}{\widetilde{N}^{\beta-2}})$ since both J and $\sqrt{\log(\widetilde{N})}$ are less than $\widetilde{N}$.

To explore the analysis above we use two illustrative simulations. For the first experiment, we generate 5000 signals with length $\widetilde{N} = NJ = 64\times64 = 4096$, using three different sparsity levels $S \in \{5, 30, 64\}$. The DFT coefficient locations are selected uniformly at random, and the corresponding nonzero coefficient values are drawn from the i.i.d. standard Gaussian distribution. Figure 7(a) plots the ratio $\frac{\Gamma_2(x,\mathbf{M})}{M}$, showing that this quantity is indeed near the upper bound of $J = 64$, indicating a favorable energy distribution. This gives support to the fact that the theoretical value of $\Gamma_2(x, \mathbf{M})$ predicted in Theorem 3.1.6 does not depend strongly on the exact value of S. For the second experiment, we fix the sparsity at $S = 5$ and vary the signal block length $J \in \{64, 200, 400\}$ (and thus the total signal length $\widetilde{N}$ changes as well). For each J we generate 5000 random signals in the same manner as before and plot in Figure 7(b) the distribution of $\frac{\Gamma_2(x,\mathbf{M})}{\widetilde{M}}$. It is clear that this value concentrates near its upper bound of 1, showing the relative accuracy of the prediction that $\frac{\Gamma_2}{M}$ scales linearly with increasing J. While some of the quantities in Theorem 3.1.6 appear pessimistic (e.g., the scaling with $\log^4(\widetilde{N})$), these simulations confirm the intuition derived from the theorem that frequency sparse signals should indeed have favorable energy distributions and therefore favorable concentration properties when measured with DBD matrices.

Because differences between frequency sparse signals are themselves sparse in the frequency domain, it follows immediately that not only do frequency sparse signals



(a) (b)

Figure 7: Histograms of the normalized quantity Γ_2 for frequency sparse signals. (a) The distribution of $\frac{\Gamma_2}{M}$ for randomly generated frequency sparse signals of length $\widetilde{N} = N \times J = 64 \times 64$ for sparsity levels $S \in \{5, 30, 64\}$. Note that $\frac{\Gamma_2}{M}$ accumulates near its upper bound of $J = 64$ for all three sparsity levels. (b) The distribution of $\frac{\Gamma_2}{\widetilde{M}}$ for randomly generated frequency sparse signals with $S = 5$ and the number of signal blocks $J \in \{64, 200, 400\}$. Note that $\frac{\Gamma_2}{\widetilde{M}}$ accumulates near its upper bound of 1.

x have favorable $\Gamma_2(x, \mathbf{M})$ values for uniform DBD matrices, but also that most differences $x_1 - x_2$ between frequency sparse signals have favorable $\Gamma_2(x_1 - x_2, \mathbf{M})$ values. Thus, when measured by a uniform DBD matrix, many families of frequency sparse signals are likely to perform favorably and predictably in the compressive signal processing scenarios outlined in Section 3.1.3.

Importantly, Theorem 3.1.6 can also allow us to guarantee the stable embedding of certain infinite collections of frequency sparse signals. In particular, for any sparse support Ω on which (22) holds uniformly, one can apply standard covering arguments (as discussed in Section 2.4) to guarantee that with a moderate total number of measurements $\widetilde{M} \gtrsim \frac{|\Omega|}{\epsilon^2} \log^4(\widetilde{N})$, a uniform DBD matrix will simultaneously approximately preserve the norm of all frequency signals supported on Ω. This fact allows one to consider compressive-domain interference cancellation (as discussed in Section 3.1.3.1 and in [43]) from a set of frequency sparse signals, where the set of possible interferers live in a known subspace of frequency sparse signals.

3.1.4.3 Difference Signals

In applications such as classification, we require a stable embedding of difference vectors between signals in a certain signal class. It is interesting to determine what signal families in addition to frequency sparse signals will give rise to difference signals that have favorable values of Γ_2 (for uniform DBD matrices) or Λ_2 (for RBD matrices).

We provide a partial answer to this question by briefly exemplifying a signal family that is favorable for measurement via uniform DBD matrices. Consider a set $Q \subset \mathbb{R}^{\widetilde{N}}$ of signals that, when partitioned into J blocks of length N, satisfy both of the following properties: (i) each $x \in Q$ has uniform energy across the J blocks, i.e., $\|x_1\|_2^2 = \|x_2\|_2^2 = \cdots = \|x_J\|_2^2 = \frac{1}{J}\|x\|_2^2$, and (ii) each $x \in Q$ has highly correlated blocks, i.e., for some $a \in (0, 1]$, $\langle x_i, x_j \rangle \geq a\frac{1}{J}\|x\|_2^2$ for all $i, j \in [J]$. The first of these conditions ensures that if $\mathbf{M} = \mathrm{diag}(M, M, \ldots, M)$, then each $x \in Q$ will have $\Gamma_2(x, \mathbf{M}) = \widetilde{M}$ and thus be highly amenable to measurement by a uniform DBD matrix. The second condition, when taken in conjunction with the first, ensures that all difference vectors of the form $x - y$ where $x, y \in Q$ will also be highly amenable to measurement by a uniform DBD matrix. In particular, for any $i, j \in [J]$, one can show that

$$\left| \|x_i - y_i\|_2^2 - \|x_j - y_j\|_2^2 \right| \leq \frac{4\sqrt{2}\|x\|_2\|y\|_2\sqrt{1-a}}{J},$$

meaning that the energy differences in each block of the difference signals can themselves have small differences. One implication of this bound is that as $a \to 1$, $\Gamma_2(x - y, \mathbf{M}) \to \widetilde{M}$.

Signal families of the form specified above—with uniform energy blocks and high inter-block correlations—may generally arise as the result of observing some phenomenon that varies slowly as a function of time or of sensor position. As an empirical demonstration, let us consider a small database of eight real-world video signals frequently used as benchmarks in the video compression community, where we will

treat each video frame as a signal block.[5] We truncate each video to have $J = 150$ frames, each of size $N = 176 \times 144 = 25344$ pixels, and we normalize each video (not each frame) to have unit energy. Because the test videos are real-world signals, they do not have perfectly uniform energy distribution across the frames, but we do observe that most frame energies are concentrated around $\frac{1}{J} = 0.00667$.

For each video, we present in Table 1 the minimum and maximum inner products $\langle x_i, x_j \rangle$ over all $i \neq j$, and we also list the quantity $\frac{\Gamma_2(x,\mathbf{M})}{M}$ for each video. As we can see, the minimum inner product for each video is indeed quite close to 0.00667, suggesting from the arguments above that the pairwise differences between various videos are likely to have highly uniform energy distributions. To verify this, we compute the quantity $\frac{\Gamma_2}{M}$ for all possible $\binom{8}{2}$ pairwise difference signals. As we are limited in space, we present in Figure 8 plots of the energies $\|x_j\|_2^2$, $\|y_j\|_2^2$, and $\|x_j - y_j\|_2^2$ as a function of the frame index j for the pairs of videos x, y that give the best (highest) and the worst (smallest) values of $\frac{\Gamma_2(x-y,\mathbf{M})}{M}$. We see that even the smallest $\frac{\Gamma_2}{M}$ value is quite close to the best possible value of $\frac{\Gamma_2}{M} = 150$. All of this suggests that the information required to discriminate or classify various signals within a family such as a video database may be well preserved in a small number of random measurements collected by a uniform DBD matrix.

3.1.4.4 Random Signals

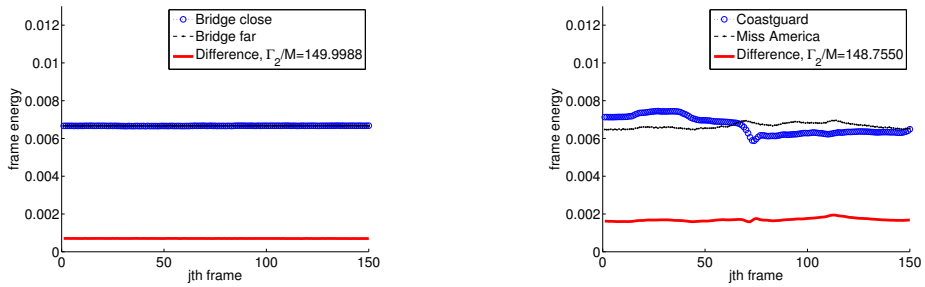
Our discussions above have demonstrated that favorable Γ_2 values (for uniform DBD matrices) and Λ_2 values (for RBD matrices) can arise for signals in a variety of practical scenarios. This is no accident. Indeed, as a blanket statement, it is true that a large majority of all signals $x \in \mathbb{R}^{\widetilde{N}}$, when partitioned into a sufficiently small number of blocks J and measured uniformly, will have favorable values of both Γ_2 and Λ_2. One way of formalizing this fact is with a probabilistic treatment such as

[5] Videos were obtained from `http://trace.eas.asu.edu/yuv/`.

Video name	Akiyo	Bridge close	Bridge far	Carphone
$\max\langle x_i, x_j\rangle$	0.00682	0.00668	0.00668	0.00684
$\min\langle x_i, x_j\rangle$	0.00655	0.00664	0.00665	0.00598
Γ_2/M	149.9844	149.9998	149.9999	149.9287

Video name	Claire	Coastguard	Foreman	Miss America
$\max\langle x_i, x_j\rangle$	0.00690	0.00742	0.00690	0.00695
$\min\langle x_i, x_j\rangle$	0.00650	0.00562	0.00624	0.00606
Γ_2/M	149.9782	149.2561	149.9329	149.9301

Table 1: The maximum and minimum inner products between all pairs of distinct frames in each video, and the quantity Γ_2/M for each video. The best possible value of Γ_2/M is $J = 150$.



(a) (b)

Figure 8: Plots of the energy distributions of individual videos and of their differences for the best video pair and the worst video pair among all possible $\binom{8}{2}$ possible video pairs. (a) The difference of the video pair, "Bridge close" and "Bridge far", giving the best value of $\Gamma_2(x-y,\mathbf{M})/M = 149.9988$. (b) The difference of the video pair, "Coastguard" and "Miss America", giving the worst value of $\Gamma_2(x-y,\mathbf{M})/M = 148.7550$.

that given in the following lemma.

Lemma 3.1.2. *Let ϕ denote a subgaussian random variable with mean 0, variance σ^2, and subgaussian norm $\|\phi\|_{\psi_2}$, and suppose $x \in \mathbb{R}^{\tilde{N}}$ is populated with i.i.d. realizations of ϕ. Let $\mathbf{M} = \mathrm{diag}(M, M, \ldots, M)$ with M fixed. Pick $\epsilon \leq \frac{\|\phi\|_{\psi_2}^2}{C_2}$ and suppose that $J \leq \frac{C_1 C_2^2 N \epsilon^2}{2\|\phi\|_{\psi_2}^4 \log(12/\epsilon)}$, where C_1, C_2 are absolute constants as given in Theorem 3.1.1. Then, with probability at least $1 - 2\exp\left(-\frac{1}{2}\frac{C_1 C_2^2 N \epsilon^2}{\|\phi\|_{\psi_2}^4}\right)$, we have*

$$\Gamma_2(x, \mathbf{M}) \geq \Lambda_2(x, M) \geq M + M\left(\frac{1-\epsilon}{1+\epsilon}\right)^2 (J-1).$$

Proof. See Appendix C.4. □

We see from Lemma 3.1.2 that when random vectors are partitioned into a sufficiently small number of blocks, these signals will have their norms preserved giving rise to $\Gamma_2(x, \mathbf{M})$ and $\Lambda_2(x, M)$ values close to their optimal value of $\widetilde{M}$. To give some illustrative numbers, numerical simulations showed that over 10000 random draws of Gaussian i.i.d. signals with $J = 16$ and $N = 64$ the average value of $\Gamma_2(x, \mathbf{M})/M$ was 15.5 and the average value of $\Lambda_2(x, M)/M$ was 12.6, which are both large fractions of the maximum possible value of 16. We also note that by using the same argument we can show that the differences of random signals will exhibit large Γ_2 and Λ_2 values. One possible use of this lemma could be in studying the robustness of block diagonal measurement systems to noise in the signal. The lemma above tells us that when restrictions are met on the number of blocks, random noise will tend to yield blocks that are nearly orthogonal and have highly uniform energies, thereby guaranteeing that they will not have their energy amplified by the matrix.

3.2 The RIP Analysis

While concentration of measure inequalities (discussed in detail in Section 3.1) are useful for applications concerning compressive signal processing, it is not evident how such results can be extended to give the RIP of block diagonal matrices with an optimum measurement rate. Obtaining the RIP for measurement matrices is advantageous as it entails signal recovery guarantees from CS. By using the ideas and techniques provided in [83, 112, 118], we shall show in this section that whenever the total number of measurements $\widetilde{M}$ scales linearly with both the sparsity of the signal and a coherence factor of the sparsifying basis, and poly-logarithmically with the ambient dimension $\widetilde{N}$, DBD and RBD matrices can indeed satisfy the RIP. We will also show that certain sparse matrices and random convolution systems considered in the CS literature can be studied in the framework of block diagonal matrices.

3.2.1 Distinct Block Diagonal Matrices

The measurement rate required for random block diagonal matrices to satisfy the RIP is dependent on the sparsity basis of the measured signals, and this dependence is reflected by a coherence factor. For DBD matrices, the coherence of a basis $\Psi \in \mathbb{C}^{\widetilde{N}\times\widetilde{N}}$ is defined as [29]

$$\mu(\Psi) := \sqrt{\widetilde{N}} \max_{p,q\in[\widetilde{N}]} |\Psi(p,q)|,$$

where $\Psi(p,q)$ is the (p,q)-th entry of Ψ. If $\{\psi_{\widetilde{n}}\}$ and $\{e_{\widetilde{n}}\}$ for $\widetilde{n} \leq \widetilde{N}$ denote the columns of Ψ and the canonical basis for $\mathbb{C}^{\widetilde{N}}$ respectively, one can easily verify that

$$\mu(\Psi) = \sqrt{\widetilde{N}} \max_{p,q\in[\widetilde{N}]} |\langle\psi_p, e_q\rangle|. \tag{23}$$

The above equation allows us to interpret $\mu(\Psi)$ as the similarity of Ψ with the canonical basis. Just like in [29], we shall see that the smaller this coherence factor $\mu(\cdot)$, the lower the required RIP measurement rate. It is easily observed that

$$1 \leq \mu(\Psi) \leq \sqrt{\widetilde{N}}.$$

The upper bound is achieved for example for the canonical basis in $\mathbb{C}^{\widetilde{N}}$ (i.e., $\mu\left(I_{\widetilde{N}}\right) = \sqrt{\widetilde{N}}$). On the other hand, the lower bound is achieved for any basis that is maximally incoherent with the canonical basis. For example, we have $\mu\left(F_{\widetilde{N}}\right) = 1$ where $F_{\widetilde{N}}$ denotes the Fourier basis in $\mathbb{C}^{\widetilde{N}}$. It turns out that a randomly chosen orthonormal basis is also highly incoherent with the canonical basis with high probability.

Lemma 3.2.1. *Let $\Psi \in \mathbb{R}^{\widetilde{N}\times\widetilde{N}}$ be a basis chosen uniformly at random from the orthogonal group $O(\widetilde{N},\mathbb{C})$.*[6] *Then, for a fixed $t \gtrsim 1$ and $\widetilde{N} \gtrsim t^2 \log\widetilde{N}$, we have*

$$\mathbb{P}\left\{\mu(\Psi) > t\sqrt{\log\widetilde{N}}\right\} \lesssim \widetilde{N}^{-t}.$$

[6] *This random basis $\Psi = [\psi_1|\cdots|\psi_{\widetilde{N}}]$ can be constructed algorithmically as follows. First, choose ψ_1 uniformly at random from the unit sphere in $\mathbb{R}^{\widetilde{N}}$. For every other $\widetilde{n} \in [\widetilde{N}] - \{1\}$ in consecutive order, $\psi_{\widetilde{n}}$ is chosen at random from the uniform distribution on the unit sphere but in the orthogonal complement of the first $\widetilde{n}-1$ columns.*

Proof. See Appendix C.5. □

Specifically, this lemma says that a randomly constructed basis will have a coherence of $\sqrt{\log \widetilde{N}}$, which is small compared to the maximum value of $\sqrt{\widetilde{N}}$.

With the coherence factor thus defined, the following theorem describes the conditions under which a DBD matrix can satisfy the RIP.

Theorem 3.2.1. *Let Φ be a $\widetilde{M} \times \widetilde{N}$ DBD matrix (12) whose non-zero entries are i.i.d., zero-mean, subgaussian random variables with subgaussian norm $\frac{\tau}{\sqrt{M}}$. Choose a RIP conditioning $0 < \delta < 1$. Let Ψ denote a basis for $\mathbb{C}^{\widetilde{N}}$, and define $\widetilde{\mu}^2(\Psi) := \min\{\mu^2(\Psi),\ J\}$. If*

$$\widetilde{M} \gtrsim_\tau \frac{S}{\delta^2} \widetilde{\mu}^2(\Psi) \log^2 S \log^2 \widetilde{N},$$

then $\Phi\Psi$ satisfies the RIP-(S, δ) except with probability $\lesssim \widetilde{N}^{-\log \widetilde{N} \log^2 S}$.

Proof. See Appendix C.7. □

To say that $\Phi\Psi$ satisfies the RIP means that Φ preserves the norms of signals that are sparse in basis Ψ, i.e., for all $x \in \mathbb{C}^{\widetilde{N}}$ such that $\|B^H x\|_0 \leq S$,

$$(1 - \delta)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta)\|x\|_2^2.$$

Now, the number of measurements required for RIP is linear in the sparsity level S and the modified coherence factor squared $\widetilde{\mu}^2(\Psi)$, and (poly-)logarithmic in the ambient dimension $\widetilde{N}$. For a well-chosen basis, this measurement rate compares favorably with that for a dense random Gaussian matrix (to within logarithmic factors in $\widetilde{N}$). For example, when measuring signals that are sparse in the frequency domain, a DBD matrix will perform nearly as well as a dense Gaussian matrix of approximately the same size.

On the other hand, when the basis Ψ is highly coherent with the canonical basis, the required number of measurements $\widetilde{M}$ to satisfy the RIP will be proportional to SJ (instead of S). While this measurement rate is J times worse than what is achievable

for signals sparse in an incoherent basis (e.g., Fourier), it is actually optimal and can even be parsimonious in certain situations. To see why it is optimal, consider the canonical basis where $\mu\left(I_{\widetilde{N}}\right) = \sqrt{\widetilde{N}}$ (thus very coherent), and consider S-sparse signals (with $S < N$) whose nonzero entries are all located within the first block of length N. Then, the energy preservation of these signals (which is a subset of all S-sparse signals) purely depends on ensuring the RIP for Φ_1. In turn, because Φ_1 is a dense sub-Gaussian matrix and thus $M \gtrsim \frac{S}{\delta^2}\log\left(\frac{N}{S}\right)$ measurements are needed for it to achieve the RIP (see Theorem 2.4.2). Therefore, it follows that the RIP for Φ is only achieved when total number of measurements obeys $\widetilde{M} \gtrsim \frac{JS}{\delta^2}\log\left(\frac{N}{S}\right)$. The required number of measurements in this case can still be parsimonious (i.e., $\widetilde{M} \ll \widetilde{N}$) if the sparsity level S of the signal is much less than the length of each signal block N.

In a sense, the discussions above highlight how the signal-dependent nature of our concentration of measure inequalities shown in Section 3.1 carries over to our RIP analysis. Additionally, if a sparsity basis Ψ was picked at random, then by combining Lemma 3.2.1 and Theorem 3.2.1, we can conclude that a measurement rate $\widetilde{M}$ proportional to $S\log^2 S\log^3 \widetilde{N}$ suffice to guarantee that a random DBD matrix will satisfy the RIP for signals sparse in Ψ with high probability. We further remark that an arbitrary permutation of the columns of a DBD matrix does not change its RIP result. The resulting permuted matrix is a sparse random matrix and is of potential interest in its own right [19].

3.2.2 Repeated Block Diagonal Matrices

To show the RIP result for RBD matrices, we require another coherence property of a basis. For this, let $x \in \mathbb{C}^{\widetilde{N}}$ be a signal and let $\alpha \in \mathbb{C}^{\widetilde{N}}$ be its representation in a basis Ψ. Therefore, $x = x(\alpha, \Psi) = \Psi\alpha$. Let $x_j = x_j(\alpha, \Psi) \in \mathbb{C}^N$ for $j \in [J]$ be the

individual blocks of the signal such that

$$x(\alpha, \Psi) = \begin{bmatrix} x_1(\alpha, \Psi) \\ \vdots \\ x_J(\alpha, \Psi) \end{bmatrix}.$$

Also define $\{\Psi_j\} \in \mathbb{C}^{N \times \widetilde{N}}$ for $j \in [J]$ be such that

$$\Psi = \begin{bmatrix} \Psi_1 \\ \vdots \\ \Psi_J \end{bmatrix}.$$

With the above definitions, we have $x_j(\alpha, \Psi) := \Psi_j \alpha$ for every j. Define $X_R(\alpha, \Psi) \in \mathbb{C}^{N \times J}$ as[7]

$$X_R(\alpha, \Psi) := \left[\begin{array}{c|c|c} x_1(\alpha, \Psi) & \cdots & x_J(\alpha, \Psi) \end{array} \right] = \left[\begin{array}{c|c|c} \Psi_1 \alpha & \cdots & \Psi_J \alpha \end{array} \right]. \tag{24}$$

Finally, define the *block-coherence* of Ψ as

$$\gamma(\Psi) := \sqrt{J} \max_{\widetilde{n} \in [\widetilde{N}]} \|X_R(e_{\widetilde{n}}, \Psi)\|_2 . \tag{25}$$

In words, $\gamma(\Psi)$ is the maximal spectral norm of the columns of Ψ when reshaped into $N \times J$ matrices. In analogy with (23), since $X_R(e_{\widetilde{n}}, \Psi)$ is a function of the columns of the canonical basis and the basis of interest Ψ, one can also think of (25) as a coherence measure between Ψ and $I_{\widetilde{N}}$ (although it would not be commutative since $\max_{\widetilde{n} \in [\widetilde{N}]} \|X_R(e_{\widetilde{n}}, \Psi)\|_2 \neq \max_{\widetilde{n} \in [\widetilde{N}]} \|X_R(\psi_{\widetilde{n}}, I)\|_2$). Quantitatively speaking, $\gamma(\Psi)$ measures the orthogonality and energy distribution between the row-blocks of every column of Ψ. If the row blocks of every column of Ψ have equal energy and are mutually orthogonal, then $\gamma(\Psi)$ will be small. In analogy with the RIP results for DBD matrices, a small $\gamma(\Psi)$ will imply a better measurement rate.

[7] This differs from the X quantity defined in (15) for the concentration of measure analysis of RBD operators by a matrix transpose.

To gain intuition on this block-coherence factor $\gamma(\Psi)$, let us compute its value for some bases. First, since every column of Ψ has unit ℓ_2-norm, it is easily observed that

$$1 \leq \gamma(\Psi) \leq \sqrt{J}.$$

Now, consider the canonical basis in $\mathbb{C}^{\widetilde{N}}$. For every $\widetilde{n} \in [\widetilde{N}]$, $X_R(e_{\widetilde{n}}, I_{\widetilde{N}})$ has a single non-zero entry having value 1, and thus $\|X_R(e_{\widetilde{n}}, I_{\widetilde{N}})\|_2 = 1$. Hence, $\gamma\left(I_{\widetilde{N}}\right) = \sqrt{J}$ which means that the canonical basis has an unfavorable block-coherence. Next, consider the Fourier basis in $\mathbb{C}^{\widetilde{N}}$, $F_{\widetilde{N}}$. The entries of the first column of $F_{\widetilde{N}}$ are all equal to $\widetilde{N}^{1/2}$. As a result, the entries of $X_R(e_1, F_{\widetilde{N}})$ are all equal to $\widetilde{N}^{1/2}$. It follows that $\|X_R(e_1, F_{\widetilde{N}})\|_2 = 1$ and therefore $\gamma\left(F_{\widetilde{N}}\right) \geq \sqrt{J}$. Again, this means that the Fourier basis has also an unfavorable block-coherence.

Finally, consider a random basis $\Psi = [\psi_1|\cdots|\psi_{\widetilde{N}}]$ constructed as in Lemma 3.2.1. Then, the columns of $X_R(e_1, \Psi)$ are J random vectors in $\mathbb{R}^N$ that are weakly dependent because $\|\psi_1\|_2 = 1$. With high probability, the length of each of the J vector is approximately $1/\sqrt{J}$ (so that $\|\psi_1\|_2 = 1$). If $J \ll N$, then these vectors are also orthogonal to one another with high probability. Thus, $\|X_R(e_1, \Psi)\|_2 \approx 1/\sqrt{J}$. Now, since the columns of Ψ have the same distribution, we have $\|X_R(e_{\widetilde{n}}, \Psi)\|_2 \approx 1/\sqrt{J}$ for every $\widetilde{n} \in [\widetilde{N}]$. Therefore, $\gamma(\Psi) \approx 1$, which is much smaller than the block-coherence of the canonical and Fourier bases. The next lemma formalizes this discussion.

Lemma 3.2.2. *Consider a random orthonormal basis Ψ constructed as in Lemma 3.2.1. For a fixed $t \leq 1$, if $J \leq N$ and $N \gtrsim t^{-2}$, then*

$$\mathbb{P}\left\{\gamma(\Psi) \gtrsim 1 + \sqrt{\frac{J}{N}} + t\right\} \lesssim \widetilde{N}^{-1}.$$

Proof. See Appendix C.6. □

This lemma tells us that a randomly drawn basis can have a block-coherence that is bounded by a constant with high probability. Thus even though the canonical and

Fourier basis have bad block-coherence, a large class of bases have block-coherence that are small.

With the block-coherence defined, we present our result on the RIP of RBD matrices.

Theorem 3.2.2. *Let Φ be a $\widetilde{M} \times \widetilde{N}$ RBD matrix* (12) *whose repeating random diagonal blocks consist of i.i.d., zero-mean, subgaussian random variable with subgaussian norm $\frac{\tau}{\sqrt{M}}$. Let Ψ denote an orthonormal basis for $\mathbb{C}^{\widetilde{N}}$, and choose a RIP conditioning $0 < \delta < 1$. If*

$$\widetilde{M} \gtrsim_\tau C \frac{S}{\delta^2} \gamma^2(\Psi) \log^2 S \log^2 \widetilde{N},$$

then $\Phi\Psi$ satisfies the RIP-(S, δ) except with a probability of $\lesssim \widetilde{N}^{-\log \widetilde{N} \log^2 S}$.

Proof. See Appendix C.7. □

Again, the theorem guarantees the RIP for $\Phi\Psi$ which means that the RBD matrix Φ preserves norms for signals sparse in the basis Ψ. The required measurement rate scales linearly with the sparsity S and the block-coherence factor $\gamma(\Psi)$, as well as logarithmically with the ambient dimension $\widetilde{N}$. This means that for bases Ψ whose $\gamma(\Psi)$ is small (and this includes the class of random bases), a random RBD matrix can perform nearly as well as a dense Gaussian random matrix. In addition to random bases which we have shown in Lemma 3.2.2 to have small $\gamma(\Psi)$ with high probability, there is a deterministic basis, which we will describe in Section 3.2.3, that achieves a similar performance.

A few additional remarks are in order. Recall that $\gamma(\Psi) = \sqrt{J}$ for signals sparse in (for example) the canonical or frequency bases, and thus the measurement rate $\widetilde{M}$ in these cases scales with JS. While worse off than the best measurement rate by a factor of J, it is in fact optimal in the sense that no less measurements can ensure the RIP for these bases. For the canonical basis, the same argument for optimality carries over from the corresponding discussion for DBD matrices. For the Fourier basis, we

shall show this optimality by constructing certain classes of periodic signals in $\mathbb{C}^{\widetilde{N}}$ that would require $\widetilde{M}$ to scales with JS. Consider, the class of signals consisting of all S-sparse linear combinations of first, $(J+1)$-th, $\cdots$,$((J-1)N+1)$-th columns of $F_{\widetilde{N}}$. If a signal x belongs to this class, then by construction, the signal blocks $x_1, \cdots, x_J$ are all the same. As a result, the different blocks of the RBD matrix Φ are basically acquiring the same measurements. Therefore, we will require M to scale with S (or equivalently $\widetilde{M}$ with JS) to preserve the norms of signals in this class. Once again, the discussions above highlight the fact that the signal-dependent nature of our concentration inequalities and the dependence of the RIP on the sparsifying basis emerge as two sides of the same coin.

3.2.3 Compressive Toeplitz and Circulant Matrices

The repeated block diagonal structure provides a powerful model for other classes of structured random matrices. In this section, we show that the structure and the RIP conditions of RBD matrices can be used to derive the RIP for partial circulant or Toeplitz matrices.

First, consider a circulant matrix $C \in \mathbb{R}^{P \times P}$, defined as

$$C = \frac{1}{\sqrt{J}} \begin{pmatrix} \xi_1 & \xi_2 & \cdots & \xi_P \\ \xi_2 & \xi_3 & \cdots & \xi_1 \\ \vdots & \vdots & & \vdots \\ \xi_P & \xi_1 & \cdots & \xi_{P-1} \end{pmatrix},$$

where $\xi = [\xi_1, \cdots, \xi_P]^T$ is an i.i.d., zero-mean, subgaussian sequence with subgaussian norms τ. In words, C is constructed from a random probe signal $\xi \in \mathbb{R}^P$ that constitute its first row, and the probe signal is circularly left-shifted as we move down the rows of C. This matrix C is then used to represent the circular convolution of a signal x with the probe signal ξ where the outcome of the convolution can be written simply as Cx.

Let $R_\Omega \in \mathbb{R}^{J\times P}$ be a subsampling matrix that performs the function of choosing a subset of J rows of a $P \times P$ matrix. The subset of rows chosen is denoted by $\Omega = \{\omega_1, \cdots, \omega_J\}$ where ω_j for $j \in [J]$ are arbitrarily chosen, non-repeating integers from 0 to $P-1$.[8] The subsampled measurements obtained can be written as $\frac{1}{\sqrt{J}} R_\Omega C x$, and the resulting subsampled circulant matrix is defined as $C_\Omega := \frac{1}{\sqrt{J}} R_\Omega C \in \mathbb{R}^{J\times P}$. Note that the $\frac{1}{\sqrt{J}}$ is simply for normalizing the columns of C_Ω.

For any S-sparse signal $x \in \mathbb{C}^P$, we can write the subsampled measurements $C_\Omega x$ as the multiplication of a RBD matrix $\Phi \in \mathbb{R}^{J\times PJ}$ and an extended vector $\widehat{x} \in \mathbb{C}^{PJ}$:

$$C_\Omega x = \frac{1}{\sqrt{J}} \begin{pmatrix} \xi^T & & \\ & \ddots & \\ & & \xi^T \end{pmatrix} \begin{pmatrix} S^{\omega_1} x \\ \vdots \\ S^{\omega_J} x \end{pmatrix} =: \frac{1}{\sqrt{J}} \Phi \widehat{x}, \tag{26}$$

where S^{ω_j} is the cyclic-shift-down-by-ω_j operator acting on column vectors in $\mathbb{C}^P$ defined as

$$S^{\omega_j}([x(1), \cdots, x(P)]^T) = \left[\overbrace{x(P-\omega_j+1), \cdots, x(P)}^{\omega_j \text{ entries}}, x(1), \cdots, x(P-\omega_j) \right]^T.$$

We now claim that we can find a basis Ψ such that $\widehat{x}$ is exactly S-sparse in this basis.

To construct this basis Ψ, first construct $\Psi' := F_J \otimes I_P \in \mathbb{C}^{PJ\times PJ}$ where $\otimes$ denotes

[8] We count the rows from 0.

the Kronecker product:

$$\Psi' = \left(\begin{array}{c|c|c} \begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix} & \cdots & \begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix} \\ \hline \vdots & \ddots & \vdots \\ \hline \begin{matrix} e^{j2\pi\frac{J-1}{J}} & & \\ & \ddots & \\ & & e^{j2\pi\frac{J-1}{J}} \end{matrix} & \cdots & \begin{matrix} e^{j2\pi\frac{(J-1)^2}{J}} & & \\ & \ddots & \\ & & e^{j2\pi\frac{(J-1)^2}{J}} \end{matrix} \end{array}\right). \tag{27}$$

Looking at (27), we see a natural partitioning of Ψ' into J^2 sub-matrices of size $P \times P$, each of which we shall denote by $\Psi'_{p,q}$ for $p, q \in [J]$. To be clear,

$$\Psi'_{p,q} = \begin{pmatrix} e^{j2\pi\frac{q-1}{J}(p-1)} & & \\ & \ddots & \\ & & e^{j2\pi\frac{q-1}{J}(p-1)} \end{pmatrix}.$$

Now, for every $p, q \in [J]$, apply the shift operator S^{w_p} to every column of $\Psi'_{p,q}$ to obtain $\Psi_{p,q} := S^{w_p}\Psi'_{p,q}$. Then, form the matrix $\Psi \in \mathbb{C}^{PJ \times PJ}$ by replacing every $\Psi'_{p,q}$ with $\frac{1}{\sqrt{J}}\Psi_{p,q}$.

With this definition for Ψ, we have $\frac{1}{\sqrt{J}}\widehat{x} = \Psi\widehat{x}_e$, where $\widehat{x}_e = [x^H,\ 0,\ \cdots,\ 0]^H \in \mathbb{C}^{PJ}$. Thus by (26), we see that $C_\Omega x = \Phi\Psi\widehat{x}_e$. For an S-sparse x, it is easy to see that $\widehat{x}_e$ is also S-sparse.

To verify that Ψ is indeed a basis, we can first check that each column of Ψ is normalized by construction. Thanks to its structure, we observe that if we pick the p-th and q-th columns of Ψ such that $(p-q) \mod P \neq 0$, their non-zero entries do not overlap and thus are orthogonal. If the p-th and q-th column are such that $(p-q) \mod P = 0$, we observe that their inner product correspond to the inner product of

the $\lceil p/P \rceil$-th and $\lceil q/P \rceil$-th columns of F_J and thus is again zero. Therefore, Ψ is a basis for $\mathbb{C}^{PJ}$. Moreover, every column and row of $X_R(e_j, \Psi)$ (basically the j-th column of Ψ but arranged into a $P \times J$ matrix) has only one nonzero entry with magnitude $\frac{1}{\sqrt{J}}$ and thus $\gamma(\Psi) = 1$.

Since Ψ is a basis with $\gamma(\Psi) = 1$, we can apply Theorem 3.2.2 to obtain that if

$$J \gtrsim \frac{S}{\delta^2} \log^2 S \log^2(PJ),$$

then except with probability $\lesssim (PJ)^{-\log(PJ)\log^2 S}$, the partial circulant matrix C_Ω satisfies the RIP-(S, δ). This corresponds to the state-of-the-art RIP result for such matrices in the literature [83].

The extension of this result to the RIP of subsampled Toeplitz matrices is easy. Suppose we have a length-N signal s which we want to convolve with the sequence ξ. Using a common trick for converting (regular) convolution into circular convolution, we zero-pad the signal s into a length-P signal x with $P = 2N - 1$, thus $x = \begin{bmatrix} s^T, & 0, & \cdots, & 0 \end{bmatrix}^T$. If s is S-sparse, then so is x. Also, $\|s\|_2 = \|x\|_2$ and therefore, the previous analysis for partial circulant matrices carries over here.

3.3 Conclusion

In this chapter, we studied compressive random DBD and RBD matrices that represent measurement systems constrained to make local measurements. First, we have derived concentration of measure inequalities for these matrices. Our experimental results confirm what our theoretical bounds suggest: that the actual probability of concentration depends on the degree of alignment between the allocation of the measurements and the energy distribution (and sometimes orthogonality) of the signal blocks. However, in situations where one can optimize the measurement allocation in anticipation of certain signal characteristics or where a fixed system may be measuring certain favorable classes of signals, we have shown that the highly structured DBD and RBD matrices can provide concentration performance that is on par with the

dense i.i.d. matrices often used in CS. We have highlighted a number of compressive signal processing applications that benefit from having a stable embedding of a finite signal family, and we have presented a modified JL lemma for block diagonal matrices that reflects the number of measurements required to ensure such a stable embedding. Finally, we have surveyed a number of signal classes whose blockwise energy distribution and/or orthogonality makes them well-suited to measurement via uniform DBD matrices or via RBD matrices. Despite not leading to state-of-the-art RIP bounds, we conclude that our nonuniform concentration results can provide a valuable tool for understanding and possibly mitigating the potential pitfalls of working with highly constrained block diagonal matrices.

Using recent techniques for bounding the suprema of chaos processes, we studied the RIP of DBD and RBD matrices next. Our main results state that these matrices can indeed satisfy the RIP but with a number of measurements depending on certain properties of the sparsifying basis. This dependence of the RIP on the sparsifying basis mirrors the signal-dependent nature of our concentration of measure inequalities. In the best case, DBD and RBD matrices perform nearly as well as dense i.i.d. random matrices generally used in CS despite having many fewer nonzero entries. Moreover, we have shown that random block diagonal matrices are intimately related to the random convolution and random Toeplitz matrices considered in the literature.

As we have discussed in the introduction, block diagonal matrices are useful for modeling distributed measurement systems. It is thus of interest to specialize our results to particular distributed systems of current interest (e.g., sensor arrays and MIMO radars). This will be discussed in further details in Section 7.1.

CHAPTER IV

STABLE MANIFOLD EMBEDDINGS

We have shown in Section 2.5 that a random orthogonal projection $\Phi \in \mathbb{R}^{M \times N}$ will provide a stable embedding of a D-dimensional submanifold $\mathcal{M} \subset \mathbb{R}^N$ whenever M scales linearly in D and logarithmically in certain other parameters of the manifold. In this chapter, we demonstrate that many other more interesting structured measurement architectures such as subsampled Fourier transforms and random convolution systems can also be used to provide stable embeddings of manifolds.[1]

4.1 Stable Manifold Embeddings

Section 4.1.1 contains a statement of our main result, showing that any matrix that satisfies the RIP (i.e., provides a stable embedding for sparse signals) can be used to form a stable embedding of a manifold. Section 4.1.2 illustrates how this fact can be used to form stable manifold embeddings from several structured matrices that have been shown to satisfy the RIP.

4.1.1 Manifold Embeddings from RIP Operators

The work in this chapter is closely related to [15] and [39] which both showed stable manifold embeddings with random orthogonal projections. The proof of each of these results requires a finite covering of points carefully chosen from the manifold and a covering of the tangent planes of those points. Using the JL lemma previously described, it then can be argued that, with high probability, a random orthogonal projection will provide a stable embedding of these points. Then, various geometric

[1] This work was performed in collaboration with Michael B. Wakin and Christopher J. Rozell. Initial results of this work have been presented in [147], while the full results are in a submitted manuscript [148]. Many thanks also go to Armin Eftekhari for useful discussions about this work.

arguments allow one to conclude that the same orthogonal projection will provide a (slightly weaker) stable embedding of the entire manifold $\mathcal{M}$. In this work, we adopt the same general proof approach but replace the JL lemma for random orthoprojectors with a JL lemma for operators satisfying the RIP. The following theorem, adapted from [84], expresses this JL lemma:[2]

Theorem 4.1.1. *Fix $0 < \rho, \epsilon < 1$ and suppose there is a finite set of points $E \subset \mathbb{R}^N$. Also suppose we have a matrix $\Phi \in \mathbb{R}^{M \times N}$ satisfying the RIP of order $S \geq 40 \log\left(\frac{4|E|}{\rho}\right)$ and conditioning $\delta \leq \frac{\epsilon}{4}$. Let $\xi \in \mathbb{R}^N$ be a Rademacher sequence (i.e., a sequence of i.i.d. equi-probable ± 1 Bernoulli random variables), construct the diagonal Rademacher matrix $D_\xi := \mathrm{diag}(\xi)$, and define $\widehat{\Phi} := \Phi D_\xi \Psi$ where $\Psi \in \mathbb{C}^{N \times N}$ is any unitary matrix. Then with probability exceeding $1 - \rho$, we have for all $x \in E$, $(1 - \epsilon)\|x\|_2^2 \leq \|\widehat{\Phi} x\|_2^2 \leq (1 + \epsilon)\|x\|_2^2$.*

In words, any operator satisfying the RIP can be used to approximately preserve the norms of any orthogonal transform of the signals in a given finite point cloud when the signs of the columns of the operator are randomly chosen. We remark that if the finite point cloud E is the set of all differences between points in another finite set $\mathcal{M} \subset \mathbb{R}^N$, then a matrix $\Phi \in \mathbb{R}^{M \times N}$ satisfying the RIP of order $S \geq 40 \log\left(\frac{4|\mathcal{M}|^2}{\rho}\right)$ (and conditioning $\delta \leq \frac{\epsilon}{4}$) in Theorem 4.1.1 can provide a stable embedding of $\mathcal{M}$ with high probability when the column signs of Φ are randomized.

Our main contribution, showing that RIP operators can be used to form stable manifold embeddings, is captured in the following theorem:

Theorem 4.1.2. *Let $\mathcal{M}$ be a compact D-dimensional Riemannian submanifold of $\mathbb{R}^N$ with geodesic regularity R, volume V, and second fundamental form uniformly bounded by $\frac{1}{\tau}$. Suppose $\Phi \in \mathbb{R}^{M \times N}$ is a matrix that satisfies RIP-(S, δ), and let $D_\xi \in \mathbb{R}^{N \times N}$ be a diagonal Rademacher matrix. Denote $\widehat{\Phi} = \Phi D_\xi \Psi$, where $\Psi \in \mathbb{C}^{N \times N}$*

[2]This theorem differs from Theorem 2.4.3 on the definition of $\widehat{\Phi}$.

is any unitary matrix. Choose any conditioning $\delta_{\mathcal{M}} < 1$ *and failure probability* ρ*. If the RIP conditioning satisfies* $\delta \leq \frac{\delta_{\mathcal{M}}}{42}$ *and the order* S *of the RIP satisfies*

$$S \geq 40 \left(2D \log \left(\frac{3528R\left(\sqrt{D/2+1}\right)(N+1)^2}{\sqrt{\pi}\delta_{\mathcal{M}}^2 \tau} \right) + \cdots \right.$$
$$\left. \cdots (2D+1) \log \left(1 + \frac{21(N+1)}{\delta_{\mathcal{M}}} \right) + \log \left(\frac{8V^2}{\rho} \right) \right),$$

then with probability exceeding $1-\rho$*,* $\widehat{\Phi}$ *provides a stable embedding of* $\mathcal{M}$ *with conditioning* $\delta_{\mathcal{M}}$*.*

The proof of this theorem can be found in Appendix D. Note that the theorem statement gives a clear recipe for both creating a stable manifold embedding from an RIP operator as well as determining how many measurements are sufficient to guarantee the desired result. The main theorem statement relates the manifold properties to the required RIP order S, which can be related to the number of measurements by the original RIP proof for the operator in question (see also Section 4.1.2). We note especially that the RIP order only scales linearly with the manifold dimension D and logarithmically with the ambient dimension N. This is especially important because most interesting RIP results also have a linear relationship between the RIP order and the number of measurements. Consequently, for such RIP results, this theorem allows the creation of a manifold embedding when the number of measurements scales linearly with the manifold dimension. Once an RIP operator is generated with a sufficient number of measurements, a manifold embedding can be created by simply randomizing the column signs of the operator.

Sometimes, such as in manifold learning algorithms (e.g., ISOMAP [129]), the main interest is in preserving the intrinsic geodesic distances between points of a data set lying on a submanifold of $\mathbb{R}^N$ instead of their extrinsic Euclidean distances. Prior work [15] has shown that operators that stably embed a manifold with respect to Euclidean distances are also stable embeddings with respect to geodesic distances.

Therefore, stable embedding operators constructed according to Theorem 4.1.2 also provide geodesic stable embeddings, guaranteeing that manifold learning algorithms can be performed significantly faster in the compressed space without much degradation [73].

4.1.2 Manifold Embeddings from Structured Matrices

As mentioned above, Theorem 4.1.2 allows us to construct operators providing stable manifold embeddings from any operator that satisfies the RIP. We illustrate the implications of our result with a few notable examples below that establish stable manifold embeddings for operators with more structure than existing results on random orthoprojectors [15]. In the corollaries that follow, we assume that $\mathcal{M}$ is a compact D-dimensional Riemannian submanifold of $\mathbb{R}^N$ with second fundamental form uniformly bounded by $\frac{1}{\tau}$, volume V, and geodesic regularity R. We also assume a fixed failure probability $0 < \rho < 1$ and conditioning $0 < \delta_{\mathcal{M}} < 1$. In what follows, we denote by $C_1, C_2, \cdots$ universal constants that do not depend on the other variables in the corollaries and that differ from corollary to corollary.

To begin, we consider a generalization of Gaussian random matrices to subgaussian random matrices (including Bernoulli, etc.).

Corollary 4.1.1 (Subgaussian matrices)**.** *Suppose $\Phi \in \mathbb{R}^{M \times N}$ is a subgaussian random matrix with independent rows or columns following the construction in [136, Thm 5.65]. If*

$$M \geq \frac{C_1}{\delta_{\mathcal{M}}^2} \left(D \log \left(\frac{RN}{\tau \delta_{\mathcal{M}}} \right) + \log \left(\frac{V}{\rho} \right) \right) \log \left(\frac{N}{D} \right),$$

then with probability greater than $1 - C_2\rho$, $\widehat{\Phi} = \Phi D_\xi$ provides a stable embedding of $\mathcal{M}$ with conditioning $\delta_{\mathcal{M}}$.

The proof of this corollary follows from the fact that such subgaussian random matrices satisfy RIP-(S, δ) with high probability whenever $M \geq C_3 \frac{S}{\delta^2} \log\left(\frac{N}{S}\right)$ [13, 136].

A natural subset of subgaussian random matrices are matrices with i.i.d., symmetric, subgaussian entries of an appropriate subgaussian norm. For this subset of matrices, both Φ and ΦD_ξ have the same distribution and thus, the stable embedding for $\mathcal{M}$ can actually use just the operator Φ rather than the operator $\widehat{\Phi}$. This last observation formally proves a remark made briefly in [15] that stable manifold embeddings can also arise from random subgaussian matrices in addition to random orthoprojectors.

To include a matrix with much more structure (i.e., not having i.i.d. entries), we also consider stable manifold embeddings by subsampled Fourier matrices.

Corollary 4.1.2 (Subsampled Fourier matrices)**.** *Suppose $\Phi \in \mathbb{R}^{M\times N}$ is a subsampled Fourier matrix whose M rows are chosen uniformly at random from the $N \times N$ DFT matrix.*[3] *If*

$$M \geq \frac{C_1}{\delta_{\mathcal{M}}^2}\left(D\log\left(\frac{RN}{\tau\delta_{\mathcal{M}}}\right)+\log\left(\frac{V}{\rho}\right)\right)\log^4(N)\log(\rho^{-1})$$

then with probability greater than $1-C_2\rho$, $\widehat{\Phi}=\Phi D_\xi$ stably embeds $\mathcal{M}$ with conditioning $\delta_{\mathcal{M}}$.

The proof of this corollary comes from the fact that subsampled Fourier matrices satisfy RIP-(S,δ) with probability greater than $1-\rho$ whenever $M \geq C_3\frac{S}{\delta^2}\log^4(N)\log(\rho^{-1})$ [111, 118]. For dimensionality reduction problems where the data lies on a manifold, this result provides an efficient measurement scheme whereby the data is pre-multiplied by a Rademacher sequence and then M coefficients from the Fourier transform of the data are randomly chosen.

In a similar direction, we also consider stable manifold embeddings from random convolutions.

Corollary 4.1.3 (Partial circulant matrices)**.** *Suppose $\Phi \in \mathbb{R}^{M\times N}$ is a partial circulant matrix whose first row is made up of i.i.d. subgaussian random variables (see [83]*

[3] *In fact, this corollary works also for subsampled DTFT matrices [111].*

for a detailed construction). If N is large enough and

$$M \geq \frac{C_1}{\delta_{\mathcal{M}}^2}\left(D\log\left(\frac{RN}{\tau\delta_{\mathcal{M}}}\right)+\log\left(\frac{V}{\rho}\right)\right)\log^4(N),$$

then with probability greater than $1-C_2\rho$, $\widehat{\Phi}=\Phi D_\xi$ stably embeds $\mathcal{M}$ with conditioning $\delta_{\mathcal{M}}$.

Here, the proof follows from the fact that partial circulant matrices satisfy RIP-(S,δ) with probability greater than $1-N^{-(\log N)(\log^2 S)}$ (hence the requirement for N to be large enough) whenever $M \geq C_3\frac{S}{\delta^2}\log^4(N)$ (for $N \geq S$) [83]. This again affords us an efficient implementation of a dimensionality reduction scheme for data residing on or near a manifold. One would first pre-process the data by multiplying its entries with a pre-chosen random Rademacher sequence. Then, one would simply convolve the processed data with a separate random subgaussian sequence and arbitrarily select M samples of the convolution output.

In some situations, one may need to apply the convolution directly on the manifold-modeled data instead of using a pre-processing step (i.e., first multiplying by a diagonal Rademacher matrix). For this, consider the matrix $\widehat{\Phi} := R_\Omega F D_\xi F^H$ where $F \in \mathbb{C}^{N\times N}$ is the DFT basis and $R_\Omega \in \mathbb{R}^{M\times N}$ is a restriction operator that selects M entries of a length-N vector (or selects M rows from an $N \times N$ matrix). Now, this matrix follows our stable embedding construction as $\Phi := R_\Omega F$ is a subsampled Fourier matrix that satisfies the RIP (as long as M is large enough) and $\Psi := F^H$ is orthonormal. Conveniently, $FD_\xi F^H$ is a circular convolution matrix with D_ξ being the (normalized) Fourier transform of the probe sequence of the convolution. Thus, the matrix $\widehat{\Phi}$ represents a subsampled convolution operation that can be used to stably embed manifold-modeled data. This idea is formalized in the follow corollary.

Corollary 4.1.4 (Random convolution matrices)**.** *Let $C_\xi \in \mathbb{C}^{N\times N}$ be a random circulant matrix such that $C_\xi := FD_\xi F^H$ where D_ξ is a random diagonal Rademacher matrix and F is the DFT basis. Let $\Omega \subset \{1,2,\cdots,N\}$ with $|\Omega| = M$ be a subset*

selected uniformly at random. If

$$M \geq \frac{C_1}{\delta_{\mathcal{M}}^2}\left(D\log\left(\frac{RN}{\tau\delta_{\mathcal{M}}}\right)+\log\left(\frac{V}{\rho}\right)\right)\log^4(N)\log(\rho^{-1}),$$

then with probability greater than $1-C_2\rho$, $\widehat{\Phi} := R_\Omega C_\xi$ *stably embeds* $\mathcal{M}$ *with conditioning* $\delta_{\mathcal{M}}$.

The proof for this corollary follows quickly from the fact that subsampled DFT matrices satisfy RIP-(S,δ) with high probability whenever $M \geq C_3\frac{S}{\delta^2}\log^4(N)\log(\rho^{-1})$ [111, 118].

To address the constraint that some systems can only take localized measurements of the signal, we also consider operators represented by a DBD matrix $\Phi \in \mathbb{R}^{MJ\times NJ}$ that is non-zero only on the diagonal blocks,

$$\Phi = \begin{pmatrix} \Phi_1 & & \\ & \ddots & \\ & & \Phi_J \end{pmatrix}.$$

The blocks $\Phi_j \in \mathbb{R}^{M\times N}$ on the diagonal consist of i.i.d. subgaussian random variables (that are also independent across the blocks). The following corollary establishes how such matrices can be used to stably embed manifold-modeled data.

Corollary 4.1.5 (DBD matrices)**.** *Let* $\Phi \in \mathbb{R}^{MJ\times NJ}$ *be a DBD matrix described above, and let* $C_\xi \in \mathbb{C}^{NJ\times NJ}$ *be the circulant matrix as described in Corollary 4.1.4. If* NJ *is large enough and*

$$MJ \geq \frac{C_1}{\delta_{\mathcal{M}}^2}\left(D\log\left(\frac{RNJ}{\tau\delta_{\mathcal{M}}}\right)+\log\left(\frac{V}{\rho}\right)\right)\log^6(NJ),$$

then with probability greater than $1-C_2\rho$, $\widehat{\Phi} = \Phi C_\xi$ *stably embeds* $\mathcal{M}$ *with conditioning* $\delta_{\mathcal{M}}$.

The proof of this corollary follows quickly from the fact that a DBD matrix Φ satisfies RIP-(S,δ) with probability exceeding $1-C_3(NJ)^{-1}$ (hence the requirement

that NJ is large enough) for frequency sparse signals (i.e., ΦF satisfies RIP) whenever $MJ \geq C_4 S \log^6(NJ)$ [143]. This corollary states that if we pre-process the data by convolving it with a random Rademacher probe, then a block-diagonal matrix (having significantly many more zeros than non-zeros) can stably embed a manifold.

As a last example, the following corollary indicates how one might be able to use a deterministic matrix construction to stably embed manifold-modeled data.

Corollary 4.1.6 (Deterministic binary matrices)**.** *Suppose* $\Phi \in \{0,1\}^{M \times N}$ *is a deterministic matrix following the construction given in [49]. If*

$$M \geq \frac{C_1}{\delta_{\mathcal{M}}^2} \left(D \log\left(\frac{RN}{\tau \delta_{\mathcal{M}}}\right) + \log\left(\frac{V}{\rho}\right) \right)^2 \log^2(N),$$

then with probability greater than $1-\rho$, $\widehat{\Phi} = \Phi D_\xi$ *provides a stable embedding of* $\mathcal{M}$ *with conditioning* $\delta_{\mathcal{M}}$.

Again, this corollary follows from the fact [49] that such matrices satisfy RIP-(S,δ) whenever $M \geq C_2 \frac{S^2}{\delta^2} \log^2(N)$. Despite the additional number of required measurements, deterministic matrices can be of interest to the CS community as it is an NP-hard problem to verify whether a randomly constructed matrix satisfies the RIP [130].

4.2 Discussions

In this chapter, we showed that all measurement operators Φ satisfying the RIP can be used to obtain a stable embedding of a manifold. Moreover, we used this main result to demonstrate several specific examples of stable manifold embeddings that represent efficient dimensionality reduction schemes and operators that model constraints on the measurement process. These include subsampled Fourier matrices, random convolution matrices, block diagonal matrices, and deterministically constructed matrices. For each of these operators, we also provided the requisite number of measurements sufficient to ensure a stable embedding of the manifold with high probability and with

a pre-determined conditioning. This result represents a combination of two directions of recent interest in the CS community: structured measurement matrices and the development of low-dimensional signal models beyond the canonical sparsity model.

While our main theorem provides a general way to construct manifold embeddings from RIP operators by paying reasonable penalties in the number of measurements, there is room for this result to be improved. Specifically, Theorem 4.1.2 could be strengthened by removing the logarithmic dependence on the ambient dimension N from the required RIP order S. This reduction by a factor of $\log(N)$ would come at the cost of the proof requiring much more sophisticated machinery involving chaining arguments as described in [39, Lemma 3.1].[4] We have chosen to present the current result using a simpler proof technique because even with the improvement described above, the final result would still require a number of measurements that depends on $\log(N)$ due to this factor arising in the RIP requirements for known matrices (as demonstrated in the corollaries of Section 4.1.2). Therefore, while this more complex proof technique could reduce the dependence on $\log(N)$, it could not entirely remove this dependence on the ambient dimension.

Second, the same proof techniques can yield similar results for stable embedding of a finite union of *non-intersecting* submanifolds. Intuitively, we expect that the number of measurements M would increase by $\log(J)$ where J is the number of manifolds and the achievable conditioning would depend on the minimum Euclidean distances between the manifolds. However, some models of interest (including sparse signals and low-rank matrices) may lie on a finite union of *intersecting* manifolds. In such cases, additional constraints on the difference manifold (i.e., $\mathcal{M}_1 - \mathcal{M}_2$ for $\mathcal{M}_1, \mathcal{M}_2$ taken from the union) may be necessary for the same proof techniques to work. For example, the difference of two sparse vectors having disjoint support lies on

[4]The fundamental technical consideration is that the current proof technique would have to be extended to consider coverings of the manifold at all scales instead of just a single scale.

a third manifold (subspace) of twice the dimension of the original manifolds and the difference of two low-rank matrices lies on a third manifold of matrices having rank twice of that of the original matrices. In these two cases, the additional constraint required for stable embedding of unions of intersecting manifolds is that the difference manifold is embedded in another slightly higher-dimensional manifold.

CHAPTER V

STABLE TAKENS' EMBEDDING

In the last two chapters, we have considered static signals modeled using sparsity (Chapter 3) or a more general manifold model (Chapter 4). In this chapter, we will instead consider dynamic signals modeled by the states $x(t) \in \mathbb{R}^N$ of a dynamical system that evolves through an ordinary differential equation

$$\dot{x} = \Psi(x), \tag{28}$$

where $\Psi : \mathbb{R}^N \to \mathbb{R}^N$ is a smooth function representing a vector field.

As before, we will endow dynamic signals with a low-dimensional structure. In particular, we consider dynamical systems whose states are confined to a low-dimensional submanifold or dissipative systems whose states have converged to a small subset of the state space called the attractor (given initial conditions in a basin of attraction) [81, 120, 125]. Using a very loose operational definition, we say that $\mathcal{M}$ is an *attractor* of a dynamical system if it is an *indivisible*, *invariant* subset of $\mathbb{R}^N$ onto which predetermined initial states of the system converge as time goes to infinity. Examples of attractors of dynamical systems inlcude fixed points, limit cycles (subsets that are topologically equivalent to circles), general submanifolds, and fractal subsets (subsets that exhibit self-similarity at different scales and thus have fractal dimensions) called chaotic attractors. A famous example of the latter is the Lorenz attractor (see Figure 9) which is the attractor of the Lorenz equations that models convection rolls in the atmosphere [125]. In what follows, we assume that the initial state of the dynamical system already lies on the manifold or attractor $\mathcal{M}$ so that for every time t, the dynamic signal $x(t)$ lies on $\mathcal{M}$.

The goal of this chapter is to answer the question of whether measurement systems

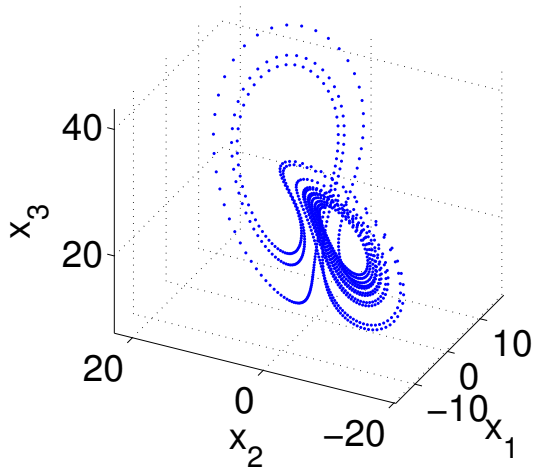


Figure 9: States of the Lorenz system are shown moving along a trajectory on the Lorenz attractor.

of dynamic signals can provide a stable embedding of attractors. In Section 5.2, we focus on linear dynamical systems (where the vector field Ψ is a matrix) and linear observation functions (where h is a vector).[1] Then, in Section 5.3, we present preliminary work on the more general case of nonlinear dynamical systems and nonlinear observation functions.[2] Before describing our work, let us first give an overview of existing results on the embeddings of such signals.

5.1 Background

5.1.1 Embeddings of Dynamical Systems

Usually through either ignorance of the variables in the system state vector $x(t)$ or limitations in sensor technologies, experimentalists only get to observe a one-dimensional time series of the dynamical system denoted by $s(t) = h(x(t))$, where $h : \mathcal{M} \to \mathbb{R}$ is a smooth function called an *observation or measurement function*. As an example,

[1] This work was performed with Christopher J. Rozell and has been published and presented in [144, 146].

[2] This work was performed in collaboration with Michael B. Wakin, Armin Eftekhari, and Christopher J. Rozell. AE and HLY contributed equally to this work, with both authors making substantial contributions to the development of the main theoretical results. A journal submission is currently being prepared for this work.

Figure 10 shows a time series formed by a measurement process that simply returns the third dimension of the system states of the Lorenz system.

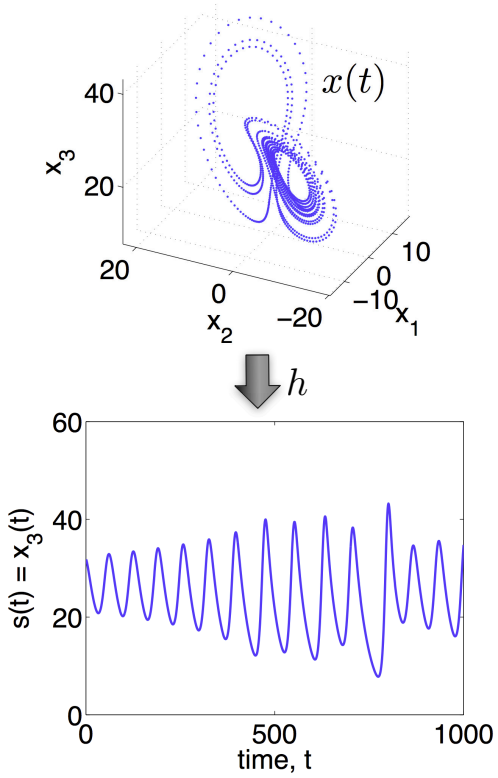


Figure 10: A time series (figure on bottom) is formed by returning the third dimension of the system states of a Lorenz system (figure on top).

Time series data has traditionally been analyzed with stochastic tools (e.g., ARMA models), where the data is typically modeled as the filtered output of white noise [24]. In the past few decades, following a seminal paper by Takens [126] coupled with research on nonlinear dynamical systems [125], a new set of geometric tools for time-series analysis has emerged. Underpinning this new set of tools is Takens' embedding theorem. Remarkably, Takens' theorem states that information about the system states can be reconstructed from sampled time-series data in spite of a severe one-dimensional compression of the states via the measurement function at every time instance.

To describe the theorem, suppose the system states reside on a submanifold $\mathcal{M} \subset$

$\mathbb{R}^N$.[3] Given a sampling time T_s, one could define the discretized dynamics of the system by computing the *flow* function $\psi_{T_s} : \mathcal{M} \to \mathcal{M}$ given by

$$\psi_{T_s}(x(t)) := \int_t^{t+T_s} \Psi(x(t))\, dt.$$

Thus, ψ_{T_s} gives us the system state at time T_s in the future, i.e.,

$$x(t + T_s) = \psi_{T_s}(x(t)). \tag{29}$$

If the sampling time T_s is fixed, we will drop the subscript T_s from ψ_{T_s} to ease notations. Then, define the *delay-coordinate map* $F_{(h,\psi)} : \mathbb{R}^N \to \mathbb{R}^M$ with M delays as

$$F_{(h,\psi)}(x(t)) := \begin{bmatrix} s(t) \\ s(t - T_s) \\ \vdots \\ s(t - (M-1)T_s) \end{bmatrix} = \begin{bmatrix} h(x(t)) \\ h\left(\psi^{-1}\left(x(t)\right)\right) \\ \vdots \\ h\left(\psi^{-M+1}\left(x(t)\right)\right) \end{bmatrix},$$

with ψ being the flow function defined in (29). In words, the delay-coordinate map with M delays is formed by stacking the M previous time-series data up into a vector. When there is no ambiguity, we will drop the subscript (h, ψ) from $F_{(h,\psi)}$.

Figure 11 illustrates the process of building this delay-coordinate map. From this figure, we see that F is a mapping from the ambient space $\mathcal{M} \subset \mathbb{R}^N$ where the system states reside to a *reconstruction space* $\mathbb{R}^M$ formed from the time series measurements. Then, Takens' embedding theorem says that for a large enough M and for a large class of measurement functions h, the delay-coordinate map F is an *embedding* of the submanifold $\mathcal{M}$:

Theorem 5.1.1. *[124, 126] Let T_s be the sampling period, $\mathcal{M}$ a K-dimensional submanifold of $\mathbb{R}^N$, and $\psi = \psi_{T_s}$ a flow on $\mathcal{M}$. Suppose that on the submanifold $\mathcal{M}$, ψ has only a finite number of periodic orbits of period less than M, and the eigenvalues*

[3] The case where the system states lie on a fractal subset will be discussed later.

of $D_p\psi$ are distinct for all p lying on each such orbit. If $M \geq 2K+1$, then there is an open and dense set of measurement functions h in the space of smooth functions $\mathcal{C}^\infty(\mathcal{M}, \mathbb{R})$ for which $F_{(h,\psi)}$ is an embedding.

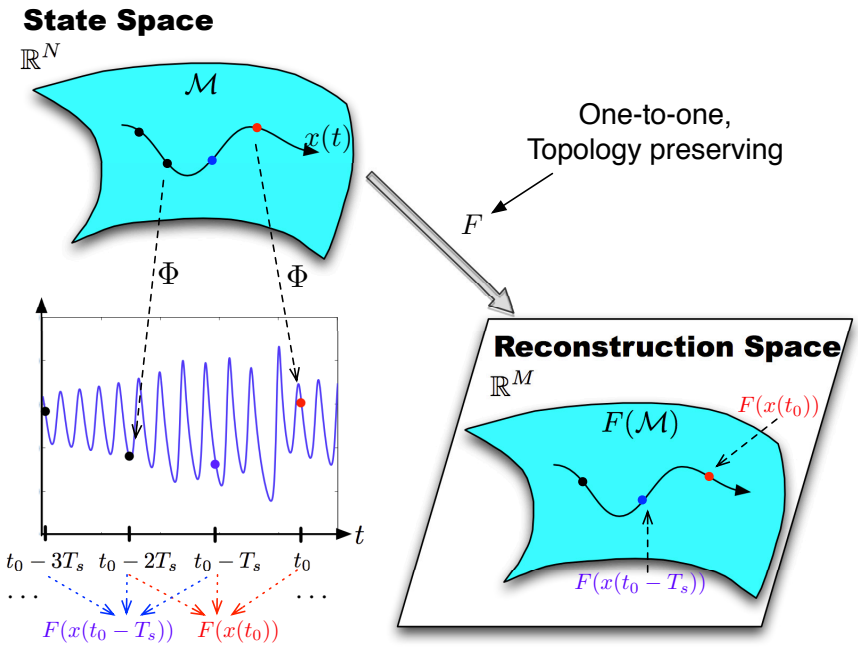


Figure 11: The embedding of a dynamical system's attractor $\mathcal{M}$ using a delay-coordinate map F is described pictorially here.

By virtue of the embedding provided by F, we can define an equivalent flow of the dynamical system in the reconstruction space as

$$\widehat{\psi} = F \circ \psi \circ F^{-1}.$$

In other words, this flow $\widehat{\psi}$ represents the same dynamical system seen under a coordinate change. Thus, characterizations of the dynamical system that do not depend on coordinates will remain unchanged in the reconstruction space. These characterizations include the number and types of fixed points [81], the correlation dimension of the attractor of the dynamical system [68], and possibly the Lyapunov exponents [59]. First, this means that even though we have no access to the hidden system states, we still can calculate these various characterizations of the dynamical system via the delay-coordinate vectors $F(x(t))$. Additionally, many time-series

algorithms make use of this topological embedding. These algorithms include time-series noise reduction [121], chaos synchronization and control [104, 108], and system identification [48, 68].

In particular, one frequently used application of Takens' theorem is time-series prediction [7, 81]. One simple method of using previous time-series data to determine future values can be described as follows [81]. Suppose we are looking to predict a time-series value at time T_s ahead of a current value $s(t) = h(x(t))$. The first step is to form a neighborhood of radius ϵ around the delay vector of the current time-series value $F(x(t))$. Denote the set of delay vectors in the neighborhood by $\mathcal{U}_\epsilon(F(x(t)))$. For each delay vector $F(x(t_k)) \in \mathcal{U}_\epsilon(F(x(t)))$ where $t_k < t$, look at its first element $s(t_k)$, and pick the next value $s(t_k + T_s)$ from the time-series data. Then, we use the average of these values $s(t_k + T_s)$ over k as a predictor of the required future time-series value.

More sophisticated time-series prediction algorithms regard future time-series values as an appropriate function of the past time-series data and attempt to learn this function from previously acquired data. Indeed, with our definition of the equivalent flow in the reconstruction space, we see that

$$\begin{pmatrix} s(t+T_s) \\ \vdots \\ s(t-(M-2)T_s) \end{pmatrix} = F(x(t+T_s)) = \widehat{\psi}(F(x(t))) = \widehat{\psi}\left(\begin{pmatrix} s(t) \\ \vdots \\ s(t-(M-1)T_s) \end{pmatrix}\right).$$

Since all but the first of the entries in $F(x(t+T_s))$ appear in $F(x(t))$, we can write our required future time-series value as

$$s(t+T_s) = \widehat{\psi}_1\left(s(t), \cdots, s(t-(M-1)T_s)\right) = \widehat{\psi}_1(F(x(t)),$$

where we see that $\widehat{\psi}_1$ is an unknown function of the past M time-series values. Then, various machine learning algorithms (e.g., artificial neural networks (ESN) [79, 90] and SVM regression [7]) can be applied to learn $\widehat{\psi}_1$ from training time-series data.

Takens' theorem has been extended by Sauer et al. [120] to include delay-coordinate embeddings of any general subset $\mathcal{M} \subset \mathbb{R}^N$. Thus, this theorem extends Takens' theorem to dissipative dynamical systems whose states converge onto a fractal subset. To quantify the dimensionality of $\mathcal{M}$ (where now there is no manifold structure), we use the concept of the box-counting dimension:

Definition 5.1.1. *[120] Consider a set $\mathcal{S} \in \mathbb{R}^N$. Suppose $\mathbb{R}^N$ is divided into cubes of size η by a grid based at points whose coordinates are η multiples of integers. Let $\mathcal{N}(\eta)$ be the number of boxes or cubes of size η that intersect $\mathcal{S}$. Then the box-counting dimension K of $\mathcal{S}$ is defined as*

$$K = \lim_{\eta \to 0} -\frac{\log \mathcal{N}(\eta)}{\log(\eta)}.$$

Note that for strange attractors, the box-counting dimension can have non-integer values. With this definition, Takens' theorem for general subsets can be stated as follows:

Theorem 5.1.2. *[120] Let T_s be the sampling period and $\psi = \psi_{T_s}$ be a (discretized) dissipative flow on a general subset $\mathcal{M} \subset \mathbb{R}^N$ of box-counting dimension K. Suppose that on $\mathcal{M}$, ψ has only a finite number of periodic orbits of period less than M, and the eigenvalues of the linearization of ψ about each such orbit are distinct. Define the finite-dimensional set $\mathcal{P} \subset \mathcal{C}^\infty(\mathcal{M}, \mathbb{R})$ as the set of all polynomials in N variables of total degree up to $2M$. If $M \geq 2K + 1$, then for any given measurement function h and for Lebesgue almost-every $\widetilde{h}$ in $\mathcal{P}$, the delay-coordinate map with M delays formed by the observation function $h + \widetilde{h}$ is an embedding of $\mathcal{M}$.*

Notice that Theorems 5.1.1 and 5.1.2 differ in their ways of describing which measurement functions will result in an embedding of $\mathcal{M}$. Theorem 5.1.1 uses the notion that any measurement function within a open and dense set of differentiable functions will work. On the other hand, Theorem 5.1.2 uses the notion that given

any measurement function h, it suffices to perturb it infinitesimally in any direction of a probe space.

5.1.2 Effects of Noise on Takens' Embedding

Under noiseless conditions, Takens' theorem ensures that the dynamics of a system under study is captured by the delay-coordinate vectors in the reconstruction space. The main problem with Takens' theorem under noise is that the choice of the sampling time T_s and the number of delays M can greatly affect the performance of algorithms that utilize delay reconstruction. Indeed, the choice of T_s and M changes the geometry of the resulting manifold in reconstruction space.

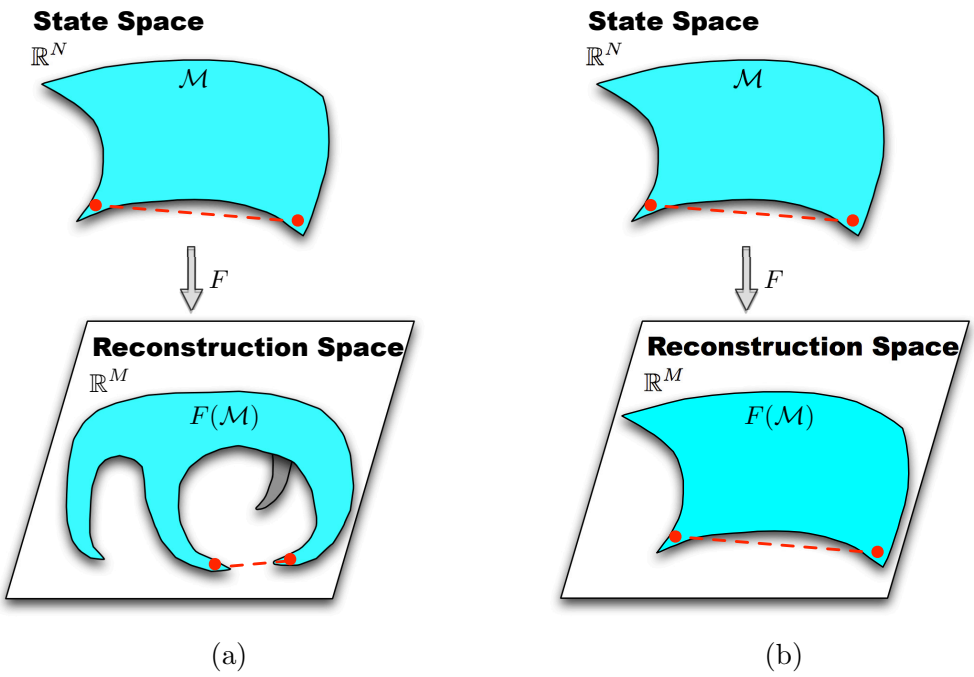
The effects of these choices on delay reconstruction is typically referred to as *redundancy* and *irrelevancy* [31, 81]. Redundancy occurs when the sampling time is too small and adjacent time-series values become very "correlated", yielding little extra "information". As such, the reconstructed manifold tends to lie stretched-out along the identity line regardless of the original manifold shape. Intuitively, taking more measurements M will tend to alleviate the flattening. However, taking more measurements with noise will increase the resulting noise power in the reconstruction space. On the other hand, irrelevancy occurs when the sampling interval is too large. In this case (especially in the presence of noise), entries in the delay-coordinate vectors can become causally disconnected. Thus, the reconstructed manifold may become unnecessarily more complex than the equivalent ambient space one.

There are many papers in the literature that give heuristic methods for choosing an "optimal" sampling time and/or number of delays. For example, in [61] Fraser and Swinney suggested using a sampling time T_s corresponding to the first minimum of the mutual information between time-series values. The resulting reconstructed manifold usually makes the quantitative and qualitative study of the dynamics easier as trajectories on the reconstructed manifold tend to be unfolded maximally to fill

the reconstruction space. Other methods of choosing T_s include looking for the first null of the autocorrelation function of the time series [81] and looking to minimize noise amplification in the reconstruction space [135]. It has also been observed that the optimal choice of M and T_s are dependent on the choice of an optimal length of time window $M \times T_s$ [31, 81, 85, 135].

The paper by Casdagli et al. [31] introduces two asymptotic quantities to describe the effects of observation noise (i.e., noise added to the time-series data), namely *noise amplification* and *distortion*. Noise amplification describes the predictive quality of the delay-coordinates vis-a-vis time-series prediction. When noise amplification is large, the prediction quality suffers. On the other hand, distortion describes the shape of an error ball in the reconstruction space when transported back into the ambient system state space. However, as these quantities are asymptotic, they usually do not inform us of how many delay coordinates M are necessary to minimize these effects. Additionally, Muldoon et al. [99] looked at a modified embedding theorem for systems corrupted by dynamical noise, considering specifically embeddings using multivariate time series outputs and taking more measurements than is typically required for a delay-coordinate map.

In our research, we propose to study the impact of noise on delay-coordinate maps in a different way. We seek to understand conditions under which delay-coordinate maps can be stable embeddings of a dynamical system for a fixed but small conditioning δ. If this can be done, the geometry of the dynamical system manifold will be preserved in the reconstruction space. As was discussed in the previous sections, this means that distances between points on the system manifold will be maintained in the reconstruction space. This favorable condition is depicted in Figure 12(b). Such stable delay-coordinate maps will ensure that small amounts of measurement noise will not result in unwanted large effects.



(a) (b)

Figure 12: This diagram shows the unstable and stable reconstructions of a dynamical system manifold $\mathcal{M}$. (a) If the delay reconstruction via delay-coordinate maps is unstable, then faraway points may be mapped close together. We see that a small addition of noise will make these points indistinguishable. (b) Here, the reconstruction is stable since distances between points are preserved.

5.2 Linear Dynamical Systems and Linear Observation Functions

In this section, we study the stable embeddings of linear systems' attractors provided by delay-coordinate maps formed from linear observation functions. For linear systems, $\Psi \in \mathbb{R}^{N\times N}$ in (28) is a matrix, and the flow ψ_T for any $T \in \mathbb{R}$ is also a matrix defined through a matrix exponential:

$$\psi_T = e^{\Psi T}.$$

Given that forming a delay-coordinate map of a specific point in the state space requires collecting samples of the system flow backward in time from that point at regular intervals T_s, we define the compact notation for the flow matrix as $\psi = \psi_{-T_s}$ so that $x(t - T_s) = \psi x(t)$. The delay-coordinate map F with M delays for the case

of linear dynamical systems and linear observation functions $h \in \mathbb{R}^N$ can then be written as a $M \times N$ matrix

$$F = \begin{pmatrix} h^T \\ h^T\psi \\ \vdots \\ h^T\psi^{M-1} \end{pmatrix}. \tag{30}$$

To ensure that the linear dynamical systems under consideration have non-trivial steady-state behavior (i.e., oscillations rather than convergence to a fixed point), we restrict our study to the class of systems $\mathcal{A}(d)$ described in the following definition.

Definition 5.2.1. *We say that a linear dynamical system in $\mathbb{R}^N$ is of* ***Class*** $\mathcal{A}(\mathbf{d})$ *for $d \leq \frac{N}{2}$ if the system matrix Ψ is real, full rank and has distinct eigenvalues. Moreover, Ψ has only d strictly imaginary*[4] *conjugate* pairs *of eigenvalues and the rest of its eigenvalues have real components strictly less than zero. The strictly imaginary conjugate pairs of eigenvalues are called the* ***$\mathcal{A}$-eigenvalues*** *and they can be expressed as $\{\pm j\theta_i\}_{i=1}^d$ where $\theta_1, \cdots, \theta_d > 0$ are d distinct numbers. The corresponding unit-norm* ***$\mathcal{A}$-eigenvectors*** *are $v_1, {v_1}^*, \cdots, v_d, {v_d}^*$. The corresponding eigenvalues of the flow matrix ψ are called the* ***$\mathcal{A}_\psi$-eigenvalues****, and are given by $\{e^{\pm j\theta_i T_s}\}_{i=1}^d$.*

Furthermore, we define

$$\Lambda := \mathrm{diag}(j\theta_1, -j\theta_1, \ldots, j\theta_d, -j\theta_d)$$

as the diagonal matrix composed of the $\mathcal{A}$-eigenvalues and

$$V := (v_1 \mid {v_1}^* \mid \cdots \mid v_d \mid {v_d}^*) \in \mathbb{C}^{N \times 2d}$$

as the concatenation of the $\mathcal{A}$-eigenvectors into a matrix with $\mathrm{rank}(V) = 2d$. Since ψ is the matrix exponential of Ψ, it is well-known that they share the same eigenvectors [98]. Therefore, if we denote $D = D_{-T_s} = e^{-\Lambda T_s}$ as the diagonal matrix comprised

[4] *A number x is* strictly imaginary *if* $\mathrm{Re}\{x\} = 0$. *This condition ensures that the system modes corresponding to these eigenvectors have persistent oscillation in the steady-state response.*

of the $\mathcal{A}_\psi$-eigenvalues, then we have $\psi V = VD$.

In order to have a meaningful notion of an embedding, the dynamical system must have its state trajectory confined to a low-dimensional attractor in the state space. Even if the system has transient characteristics from a given starting point, the embedding of a system is only considered in steady-state when these transients have disappeared. Considering the steady-state dynamics of the system, we make explicit the notion of an *attractor* (for a linear dynamical system) through the following definition.

Definition 5.2.2. *Let a linear dynamical system be of class $\mathcal{A}(d)$ and let $x_0 = V\alpha_0 \in \mathbb{R}^N$ for some $\alpha_0 \in \mathbb{C}^{2d}$ be an arbitrary initial state of the system.*[5] *We define the* ***attractor*** *of this linear dynamical system to be $\mathcal{M} = \left\{x \in \mathbb{R}^N \mid x = Ve^{\Lambda t}\alpha_0 \ , \ t \in \mathbb{R}\right\}$.*

It is easy to see that $\mathcal{M}$ lives in the span of V. Also, the attractor of the system clearly depends on the initial state of the system. Because our main result does not depend on the choice of initial state, we will simply refer to the fixed attractor as $\mathcal{M}$ and suppress the implicit dependence on the initial state. Additionally, one can check that this definition meets the fundamental notion of an attractor, i.e., that any point on the attractor $\mathcal{M}$ when projected backwards (or forward) in time by ψ will remain on $\mathcal{M}$. Specifically, for any $x \in \mathcal{M}$, we can write $x = V\alpha_x$, where $\alpha_x = e^{\Lambda t_x}\alpha_0$ for some $t_x \in \mathbb{R}$. Then we see that for some D (the diagonal matrix comprised of the $\mathcal{A}_\psi$-eigenvalues as defined earlier) and any $k \in \mathbb{Z}$,

$$\psi^k x = \psi^k V\alpha_x = VD^k\alpha_x,$$

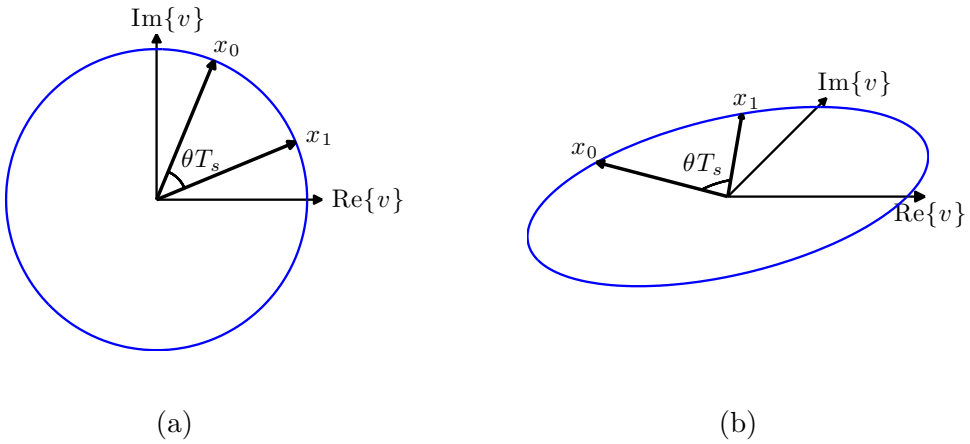
meaning that x remains on the attractor even when it is projected forward or backward in time. Finally, one can show that for each i the state $x(t)$ is moving in an elliptical orbit on the span of $\mathrm{Re}\{v_i\}$ and $\mathrm{Im}\{v_i\}$ with angular speed proportional to $\theta_i T_s$.

[5] *We only need to consider x_0 in the span of the columns of V because any orthogonal components vanish in steady-state.*

For clarity and to build intuition, we give two brief examples where $N = 2$, $d = 1$ and $T_s = 1$. For the first example, consider a dynamical system of class $\mathcal{A}(d)$ with $\mathcal{A}$-eigenvalue $\theta = \frac{\pi}{4}$ and $\mathcal{A}$-eigenvector $v = \frac{1}{\sqrt{2}}[1,\ j]^T$. Shown in Figure 13(a) is the resulting circular attractor of this system, along with the real and imaginary components of the $\mathcal{A}$-eigenvector and a pair of states separated in time by T_s (which corresponds to a separation of θT_s in angle). For the second example, consider a dynamical system of class $\mathcal{A}(d)$ with the same parameters except that the $\mathcal{A}$-eigenvector is now defined as $v = [0.8165 + 0.4082j,\ -0.4082j]^T$. Shown in Figure 13(b) is the resulting elliptical attractor and state time samples, illustrating that the angular speed is unchanged at θT_s. In both of these examples, the elongation of the ellipse is determined by the inner product between $\mathrm{Re}\{v\}$ and $\mathrm{Im}\{v\}$, which governs how well the attractors fill the dimensions of the state space that it occupies. While this is intuitive to visualize in the present case of $d = 1$, for general $d > 1$ this elongation is determined by the ratio between the smallest and largest eigenvalues of $V^H V$, denoted A_1 and A_2, respectively. When $A_1 = A_2$, the system state revolves around a circle when projected onto each of the subspaces spanned by $\mathrm{Re}\{v_p\}$ and $\mathrm{Im}\{v_p\}$ for $p = 1, \cdots, d$, and the resulting attractor is a product of these circular orbits. However when $A_2 \gg A_1$, the projection of the attractor onto some (or all) of these subspaces will be a highly elongated ellipse, therefore not equally filling the dimensions of the state space that it occupies.

5.2.1 Insufficiency of Takens' Embedding Theorem for Linear Dynamical Systems

We are interested in the question of when the one-to-one property of F described in Takens' Embedding Theorem (Theorem 5.1.1) can be improved to become a stable embedding where F is (nearly) an isometry that preserves the geometry of $\mathcal{M}$. To do so, first let us augment the definition of stable embeddings to include a scaling constant C:



(a) (b)

Figure 13: Examples of attractors of linear dynamical systems of class $\mathcal{A}(d)$ in $\mathbb{R}^N$ for $N = 2$ and $d = 1$ with sampling interval $T_s = 1$. (a) A system attractor when $\theta = \frac{\pi}{4}$ and $v = \frac{1}{\sqrt{2}}[1,\ j]^T$. This results in a circular attractor where the system progresses at an angular speed determined by θ. (b) A system attractor when $\theta = \frac{\pi}{4}$ and $v = [0.8165 + 0.4082j,\ -0.4082j]^T$. Here the system also progresses at the same angular speed, but the attractor is now an ellipse.

Definition 5.2.3. *Suppose we have a dynamical system in $\mathbb{R}^N$ that converges to an attractor $\mathcal{M}$ and a linear map $F : \mathbb{R}^N \to \mathbb{R}^M$. We say that F is a* stable embedding *of $\mathcal{M}$ with* conditioning δ *if for all $x, y \in \mathcal{M}$ and for some* scaling constant $C > 0$, *we have*

$$C(1-\delta) \leq \frac{\|F(x) - F(y)\|_2^2}{\|x - y\|_2^2} \leq C(1+\delta). \tag{31}$$

Note that smaller values of δ in the above definition imply a more stable embedding because it guarantees that the map is closer to an isometry. We also note that preservation of Euclidean distances also implies that the geodesic distances between points on the attractor are preserved [15]. Because Taken's result only tells us that the delay-coordinate map F is a one-to-one mapping, it does not guarantee any specific value of the conditioning, meaning that δ could be arbitrarily close to 1 and the embedding could be highly unstable.

To see why Takens' Embedding can be insufficient, we present an illustrative example where the conditioning of the embedding can be made arbitrarily bad when M is the minimum number of delays necessary to satisfy the sufficient conditions of

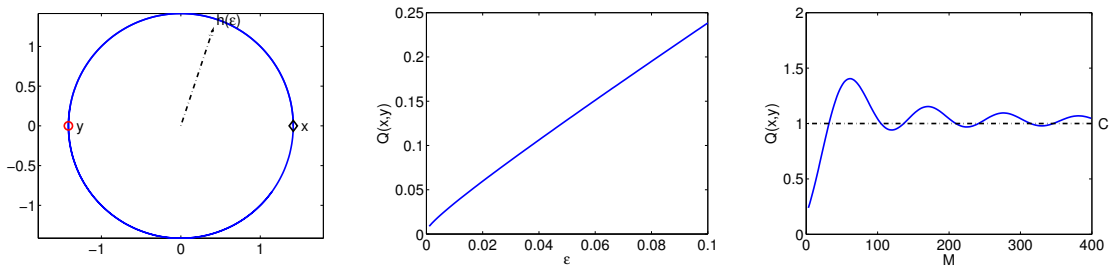
Takens' Embedding Theorem, Theorem 5.1.1. Consider a linear system of class $\mathcal{A}(1)$ with $N = 2$, $T_s = 1$, $\mathcal{A}$-eigenvalue $\theta = 0.03$ and $\mathcal{A}$-eigenvector $v = \frac{1}{\sqrt{2}}[1,\ j]^T$. This system has a circular attractor as depicted in Figure 14(a). We set the observation function to be $h = \sqrt{\frac{2}{M}}[\sqrt{\epsilon},\ \sqrt{1-\epsilon}]^T$.[6] Given a particular pair of points x, y on opposite ends of the circular attractor (shown in Figure 14(a)), we examine the ratio $Q(x,y) = \frac{\|F(x)-F(y)\|_2^2}{\|x-y\|_2^2}$, where F is the delay-coordinate map given in (30). Note that if F is a perfect isometry then $Q(x,y) = 1$, and we must have $Q(x,y) > 0$ for F to be one-to-one. Fixing the number of measurements at $M = 3$ (the minimum required by Takens' theorem), Figure 14(b) shows the behavior of $Q(x,y)$ for this pair of points as a function of ϵ. We see that while meeting the sufficient conditions of Takens' Theorem, $\lim_{\epsilon\to 0} Q(x,y) = 0$. Stated another way, by adjusting the parameter ϵ, the conditioning of F can be made arbitrarily bad for this pair of points. To see that this is not simply a bad pairing of the measurement function to the system, note that for any admissible choice of h there would exist a pair of points that would behave the same way.[7] To explore this example further, Figure 14(c) plots $Q(x,y)$ with $\epsilon = 0.1$ and varying M from 3 to 400. We see that with increasing M, the ratio $Q(x,y)$ increases, oscillates and converges to a value of $C = 1$. This provides evidence suggesting that as M increases, the conditioning of F improves because the distance between this pair of points is preserved with increasing fidelity. This effect is not predicted by Takens' Embedding Theorem, Theorem 5.1.1, but will be shown in our main results in Section 5.2.2.2.

5.2.2 Stable Embeddings for Linear Dynamical Systems

In this section, we present our main technical results. We first present a preliminary result in Section 5.2.2.1 that gives explicit sufficient conditions on the system and

[6] As will be described in Theorem 5.2.2, the observation function is normalized so that we have scaling constant of $C = 1$ regardless of M.

[7] One can imagine this by rotating the points x, y by an angle equivalent to the angle between the new measurement function and the given h.



(a) (b) (c)

Figure 14: Examining the conditioning of Takens' embeddings. (a) The large (blue) circle shows the attractor of the linear system. The (black) diamond and (red) circle markers show 2 different points x, y that we pick on the opposite ends of the attractor. The arrow depicts the measurement function $h(\epsilon)$. (b) The graph shows $Q(x, y)$ for the points x, y in Figure 14(a) over a range of values of ϵ from 0.01 to 0.1. The number of measurements M is fixed at 3, the minimum required by Takens' theorem. (c) Here $Q(x, y)$ is plotted for M ranging from 3 to 400 (with ϵ fixed at 0.1), suggesting a near isometry for F as M increases.

observation functions to guarantee that the delay-coordinate map is a one-to-one map of the state space attractor. This is akin to Takens' Embedding Theorem, and we present it here to highlight the specific differences that arise under our restrictions (linear systems and measurement functions) and when seeking explicit conditions on system and measurement pairs (as opposed to the conditions for generic observation functions in Takens' theorem). We then present our main technical contribution in Section 5.2.2.2, giving explicit conditions on the system and observation function for the delay-coordinate map to be a stable embedding of the attractor with specific guarantees on the conditioning number of the embedding.

5.2.2.1 Takens' Embeddings

The following theorem gives conditions on the system and the observation function such that the delay-coordinate map F is a one-to-one mapping. This is analogous to Takens' Embedding Theorem, Theorem 5.1.1, in the context of linear dynamical systems and linear observation functions.

Theorem 5.2.1 (Linear Takens' Embedding [144]). *Assume a linear dynamical system of class $\mathcal{A}(d)$ in $\mathbb{R}^N$ that is in steady state. Choose $T_s > 0$ to be the sampling interval, $h \in \mathbb{R}^N$ to be the observation function, and denote by F the delay-coordinate map with M delays as defined in* (30). *Suppose that $M \geq 2d$, the $\mathcal{A}_\psi$-eigenvalues $\{e^{\pm j\theta_i T_s}\}$ are distinct and strictly complex,*[8] *and $v_i^H h \neq 0$ for all $i = 1, \cdots, d$. Then, for all distinct pairs of points $x, y \in \mathcal{M}$, F satisfies* (31) *for some constants C and $\delta < 1$.*

Proof. See Appendix E.1. □

To explore the differences that arise in our specific setting of linear systems and linear observation functions, we compare the conditions of this theorem with that of Takens' theorem. First, we notice that the conditions on the measurement operation are very similar. Theorem 5.2.1 requires $M \geq 2d$, which is similar to Takens' $M > 2d$ and likely only different because of the specific structure of our attractors. There is also a close correspondence with the other condition on the measurement function $v_i^H h \neq 0$. This requirement is an explicit condition on the relationship between the system and observation function ensuring that the observation function can capture some information from every dimension of the attractor. We note that (Lebesgue) almost-every $h \in \mathbb{R}^N$ will satisfy this condition, and so we find that this is just a more explicit version of Takens' result that "almost-every" h ensures an embedding.

Next, we compare our conditions on the system with those imposed by Takens' theorem. Theorem 5.2.1 requires that the $\mathcal{A}_\psi$-eigenvalues are distinct and strictly complex, which is equivalent to having $e^{j\theta_p T_s} \neq e^{\pm j\theta_q T_s}$ (distinct) and $e^{j\theta_p T_s} \neq \pm 1$ (strictly complex) for all $p \neq q$ and $p, q = 1, \cdots, d$. While this requirement implies[9]

[8] *We say that a number x is* strictly complex *if* $\mathrm{Im}\{x\} \neq 0$.

[9] This implication can be shown by contradiction. Pick any $1 \leq k \leq 2d$ and suppose that $\mathcal{M}$ has at least a periodic orbit of Ψ with period kT_s. This would be equivalent to saying that $e^{j\theta_p kT_s} = \left(e^{j\theta_p T_s}\right)^k = 1$ for all p, meaning that for each p from 1 to d the quantity $e^{\pm j\theta_p T_s}$ is uniquely one of the k roots of unity. However this is impossible as there are $2d$ distinct and strictly

that $\mathcal{M}$ does not have periodic orbits of period kT_s for $k = 1, \cdots, 2d$ (thus satisfying Takens' condition), our condition is actually more stringent than this restriction on periodic orbits (likely due to our restricted class of linear observation functions). We note that since $\{\theta_i\}_{i=1}^d$ are distinct by definition, this condition is dependent on the choice of sampling interval T_s. One can verify that choosing $T_s < \frac{\pi}{\max\{\theta_i\}}$ is sufficient (but not necessary) to meet the condition of the theorem.

5.2.2.2 Stable Takens' Embeddings

Before presenting our main result giving conditions for a stable embedding of a dynamical system in a delay-coordinate map, it will be useful to define and understand the following quantities that characterize how well-behaved the system and measurement process are both individually and jointly. First, we define $\kappa_1 = \min_{i\in\{1,\ldots,d\}} \left\{ \frac{|v_i^H h|}{\|h\|_2} \right\}$ and $\kappa_2 = \max_{i\in\{1,\ldots,d\}} \left\{ \frac{|v_i^H h|}{\|h\|_2} \right\}$ characterizing the minimum and maximum projection of the (normalized) observation function on the $\mathcal{A}$-eigenvectors. Roughly speaking, these quantities are an indication of the disparity between the dimensions of the system attractor that are best and worst matched to the observation function. One would expect that a measurement system is most efficient when it observes all parts of the attractor equally such that $\kappa_1 \approx \kappa_2$. Second, we define A_1, A_2 as the smallest and largest eigenvalues of $V^H V$, respectively. As we discussed (at the introduction of Section 5.2), these quantities describe how well the system attractor fills the dimensions of the state space that it occupies (i.e., when $A_2 \gg A_1$, the attractor is very elongated in the state space). Again, we would expect that a system will be most amenable to observation when it fills the space such that $A_1 \approx A_2$.

Finally, we define $\nu := \max_{p\neq q} \left\{ |\sin(\theta_p T_s)|^{-1}, \left|\sin\left(\frac{(\theta_p-\theta_q)T_s}{2}\right)\right|^{-1}, \left|\sin\left(\frac{(\theta_p+\theta_q)T_s}{2}\right)\right|^{-1} \right\}$, which will also bound the constants associated with the stable embedding. Notice that the first term is large if $\theta_p T_s$ is small for some p (or that $\theta_p T_s \approx k\pi$ for some

complex values of $\{e^{\pm j\theta_p T_s}\}$ and there are only $k \leq 2d$ roots of unity (including ± 1 which are not allowed), and hence we have a contradiction.

integer k), meaning that the system state proceeds in the span of $\mathrm{Re}\{v_p\}$ and $\mathrm{Im}\{v_p\}$ at a slow pace, thus not producing much diversity in consecutive measurements of the system along these dimensions. The second term is large if $\theta_p T_s - \theta_q T_s$ is small (or near $k\pi$) for some $p \neq q$ and $p, q = 1, \cdots, d$, implying that the system state is proceeding in the subspaces spanned by $\mathrm{Re}\{v_p\}, \mathrm{Im}\{v_p\}$ and $\mathrm{Re}\{v_q\}, \mathrm{Im}\{v_q\}$ at almost the same rate. This condition would be unfavorable because the system will take an extremely long time to display enough diversity to determine that it is actually traveling on two separate subspaces instead of one. The third term is similar to the second term if we write $\theta_p T_s + \theta_q T_s = \theta_p T_s - (-\theta_q T_s)$. Thus if $\theta_p T_s \sim -\theta_q T_s$, then the system is again proceeding on two subspaces at almost the same rate (although the system is proceeding in one of the subspaces in the "opposite" direction).

Armed with these definitions, we now present our main result giving deterministic, explicit and non-asymptotic guarantees on the conditioning of the delay-coordinate map.

Theorem 5.2.2 (Stable Linear Takens' Embedding)**.** *Assume a linear dynamical system of class $\mathcal{A}(d)$ in $\mathbb{R}^N$ that is in steady state. Choose $T_s > 0$ to be the sampling interval, $h \in \mathbb{R}^N$ to be the observation function such that $\|h\|_2^2 = \frac{2d}{M}$, and denote by F the delay-coordinate map with M delays as defined in* (30)*. Suppose that $M > \left((2d-1)\,\frac{A_2\kappa_2^2}{A_1\kappa_1^2}\,\nu\right)$, the $\mathcal{A}_\psi$-eigenvalues $\{e^{\pm j\theta_i T_s}\}$ are distinct and strictly complex, and $v_i^H h \neq 0$ for all $i = 1, \cdots, d$. Then for all distinct pairs of points $x, y \in \mathcal{M}$, F satisfies* (31) *with constants $C := d\left(\frac{\kappa_1^2}{A_2} + \frac{\kappa_2^2}{A_1}\right)$ and $\delta := \delta_0 + \delta_1(M)$, where:*

$$\delta_0 := \frac{A_2\kappa_2^2 - A_1\kappa_1^2}{A_2\kappa_2^2 + A_1\kappa_1^2}, \qquad \delta_1(M) := \frac{(2d-1)\nu}{M}\left(\frac{2A_2\kappa_2^2}{A_2\kappa_2^2 + A_1\kappa_1^2}\right). \tag{32}$$

Proof. See Appendix E.1. □

We first note that the sufficient conditions of this theorem are the same as those in Theorem 5.2.1, except that the required number of measurements is larger to ensure

specific guarantees on the conditioning number δ (i.e., $\delta < 1$). Also, note that this theorem requires an observation function with a particular norm $\|h\|_2^2 = \frac{2d}{M}$. This normalization is to remove from C any dependence on the number of measurements M and the dimension of the attractor $2d$ (since κ_1^2 and κ_2^2 both scale inversely with d). The normalization plays no other significant role in the proof (and therefore could be eliminated without losing generality, but at the expense of clarity).

To understand the implications of Theorem 5.2.2, we examine the behavior of the conditioning number δ as it is the main quantity of interest. In the theorem statement, δ is a sum of δ_0 (which does not depend on M) and $\delta_1(M)$ which is positive for all M and for which $\lim_{M\to\infty} \delta_1(M) = 0$. Thus, we see that by taking more observations one could drive the conditioning guarantee for the mapping to $\delta = \delta_0$, *but not below*. In other words, some system and measurement pairs will have a plateau preventing the conditioning guarantee for the delay-coordinate map from improving beyond a fundamental limit. This is in contrast with CS results where the conditioning can be continually improved by taking more measurements. Indeed, in order to get arbitrarily good conditioning we would need $\delta_0 = 0$, which happens if and only if

$$A_2\kappa_2^2 - A_1\kappa_1^2 = 0 \;\Leftrightarrow\; \frac{A_2}{A_1} = \frac{\kappa_1^2}{\kappa_2^2} = 1.$$

Recall that $A_1 = A_2$ implies that the attractor $\mathcal{M}$ maximally fills the subspace spanned by V and $\kappa_1 = \kappa_2$ means that the observation function h projects equally onto the $\mathcal{A}$-eigenvectors. Thus even with an infinite number of measurements, the delay-coordinate map can only be guaranteed to be an exact isometry ($\delta = 0$) when the system and observation function maximally fill and measure the subspace containing the attractor.

The quantity $\delta_1(M)$ can be used to determine the number of measurements necessary to ensure that the conditioning number δ is within ϵ of the optimal value δ_0. To find the required number of measurements to meet this target $\widehat{M}(\epsilon)$, we set $\delta_1(M) = \epsilon$

and solve (32) for M to get

$$\widehat{M}(\epsilon) = \frac{(2d-1)\nu}{\epsilon}\left(\frac{2A_2\kappa_2^2}{A_2\kappa_2^2 + A_1\kappa_1^2}\right). \tag{33}$$

By multiplying the numerator and denominator by $\frac{1}{A_2\kappa_2^2}$ and noting that $0 < \frac{A_1\kappa_1^2}{A_2\kappa_2^2} \leq 1$, we can deduce that $\frac{(2d-1)\nu}{\epsilon} \leq \widehat{M}(\epsilon) < \frac{2(2d-1)\nu}{\epsilon}$. One immediate application of this fact is that we can calculate the number of measurements necessary to guarantee a stable embedding for the delay-coordinate map with a specified conditioning $\delta \in (\delta_0,\ 1)$, which is made precise in the following corollary.

Corollary 5.2.1. *Suppose we have a linear system of class $\mathcal{A}(d)$, observation function h and sampling time T_s such that the conditions of Theorem 5.2.2 are satisfied. Choose any $0 < \epsilon < (1 - \delta_0)$. If the delay-coordinate map F defined in* (30) *has a number of delays M chosen to satisfy $M \geq \frac{2(2d-1)\nu}{\epsilon}$, then F is a stable embedding of $\mathcal{M}$ with conditioning $\delta \leq \delta_0 + \epsilon$.*

The proof of this corollary is not shown, but follows immediately from Theorem 5.2.2. While the linear scaling with d seen in this result is in line with state-of-the-art CS results, we see that in contrast to typical CS results $\widehat{M}(\epsilon)$ does not depend on the ambient dimension N. Also note that $\widehat{M}(\epsilon)$ depends strongly on the $\mathcal{A}$-eigenvalues via the quantity ν. In contrast, the interactions of the $\mathcal{A}$-eigenvectors and the observation function h determine the lower bound on the conditioning δ, as evidenced by the roles played by the quantities A_1, A_2 and κ_1, κ_2 in the formula for δ_0.

5.2.3 Simulation Experiments

While the main result in Theorem 5.2.2 is encouraging, it remains to be shown that (i) the theoretical quantities actually reflect the salient embedding characteristics seen in system and measurement combinations, and (ii) having a stable embedding actually improves our ability to infer information about a hidden attractor. For example, it is important to know if the fundamental limits on the embedding quality $\delta(M)$ are

artifacts of our proof technique or are empirically observed. If these limits on the embedding quality are actually present, it is also important to know if the related bounds are tight, both in their asymptotic values and in terms of their convergence speed as M increases. Finally, for a stable embedding to be a valuable goal, we need to demonstrate that achieving this goal results in improved performance in specific tasks performed in the reconstruction space. This section will use a series of simple simulations to explore these aspects of our theoretical results.

As a general approach, each simulation in Sections 5.2.3.1 and 5.2.3.2 below involve creating an observation function h and a test system of dimension $N = 50$ in class $\mathcal{A}(d)$ (defined by $\mathcal{A}$-eigenvalues and $\mathcal{A}$-eigenvectors) so that the conditions of Theorem 5.2.2 are satisfied. We choose the arbitrary initial point x_0 defining the attractor such that $\alpha_0 = [1, \cdots, 1]^T$ and $x_0 = V\alpha_0$, and we assume a sample time of $T_s = 1$. For a single trial, we generate a random pair of points on the attractor x and y by choosing uniform random numbers t_x, t_y from $(0, 10000)$ and assigning $x = Ve^{\Lambda t_x}\alpha_0$ and $y = Ve^{\Lambda t_y}\alpha_0$. In other words, we start the system from the (arbitrary) initial condition and stop it after a random amount of time to get a single point on the attractor. We then vary M from 1 to 200, and run 1000 trials for each M (renormalizing h for each M as per Theorem 5.2.2). For each trial we calculate the quality of the conditioning $Q(x, y) = \frac{\|F(x)-F(y)\|_2^2}{\|x-y\|_2^2}$, and for each M record the largest and smallest value of $Q(x, y)$ (denoted $\max\{Q\}$ and $\min\{Q\}$, respectively) as a way to quantify how the conditioning changes with the number of measurements. In the subsequent plots the dotted lines represent $C(1 \pm \delta_0)$, showing the theoretical asymptotic bounds on the conditioning quality $Q(x, y)$, and the dashed lines are the theoretical bounds on the conditioning $C(1 \pm \delta(M))$ given by Theorem 5.2.2.

5.2.3.1 Bounds on the Embedding Quality

One of the fundamental characteristics of Theorem 5.2.2 is that in general, the bound on the embedding quality $\delta(M)$ approaches $\delta_0 \neq 0$ as M increases rather than approaching zero as is typical in CS results. The first question to ask is whether pairs of systems and observation functions can actually display such a plateau as predicted, or whether the conditioning instead continually improves with more measurements. To demonstrate this effect, we generate a simulation as described above with $d = 3$, choosing the $\mathcal{A}$-eigenvalues $\{\theta_i\}_{i=1}^{d}$ uniformly at random from $(0, \pi)$, and taking care to ensure that the resulting $\mathcal{A}_\psi$-eigenvalues are distinct and strictly complex to satisfy the conditions of Theorem 5.2.2. We then create the $\mathcal{A}$-eigenvectors by letting $v_i = \frac{1}{\sqrt{2}}(e_{2i-1} + je_{2i})$, where $\{e_i\}$ are the canonical basis vectors in $\mathbb{R}^N$. This choice of $\mathcal{A}$-eigenvectors ensures that $A_1 = A_2$. To generate a generic observation function h, we first create a vector $c \in \mathbb{R}^N$ such that $c = \sum_{i=1}^{d}((1 + w_{2i-1}) \operatorname{Re}\{v_i\} + (1 + w_{2i}) \operatorname{Im}\{v_i\})$, where the $\{w_i\}$ are i.i.d. Gaussian random variables of zero mean and variance 0.1. Thus c is a (random) linear combination of the vectors that form the subspace of the attractor. For each M we let $h = h(M) = \sqrt{\frac{2d}{M}}\frac{c}{\|c\|_2}$ so that $\|h\|_2^2 = \frac{2d}{M}$ to meet the conditions of Theorem 5.2.2. Note that the small variance of $\{w_i\}$ produces $\{|v_i^H h|^2/\|h\|_2^2\}$ centered tightly around 1, making δ_0 small (due to $A_1 = A_2$ and κ_1, κ_2 both close to 1).[10] The specific parameters in this simulation are shown in Table 2.

The results for this simulation are shown in Figure 15(a). We see from the behavior of $\max\{Q\}$ and $\min\{Q\}$ that the embedding does indeed reach a fundamental limit where the conditioning does not improve with more measurements. Furthermore, we see in this case that this plateau is correctly captured by the value $C(1 \pm \delta_0)$ as described in Theorem 5.2.2. Additionally, the bounds $C(1 \pm \delta(M))$ do contain

[10] The random variables $\{w_i\}$ are used to ensure that κ_1, κ_2 are close to, but not exactly equal to 1. The case where $\kappa_1 = \kappa_2 = 1$ is considered in the simulation in Figure 15(b).

Index i	1	2	3	4	5	6
θ_i (rad)	2.3129	0.1765	1.4861	—	—	—
$\lvert v_i^H h\rvert^2/\lVert h\rVert_2^2$	0.8346	1.1637	1.0017	—	—	—
$\lambda_i(V^H V)$	1	1	1	1	1	1

Table 2: Parameters for the simulation shown in Figure 15(a). In this case the relevant quantities are $A_1 = A_2 = 1$, $\kappa_1 = 0.8346$, $\kappa_2 = 1.1637$, $\nu = 5.6954$ and $\delta_0 = 0.1647$.

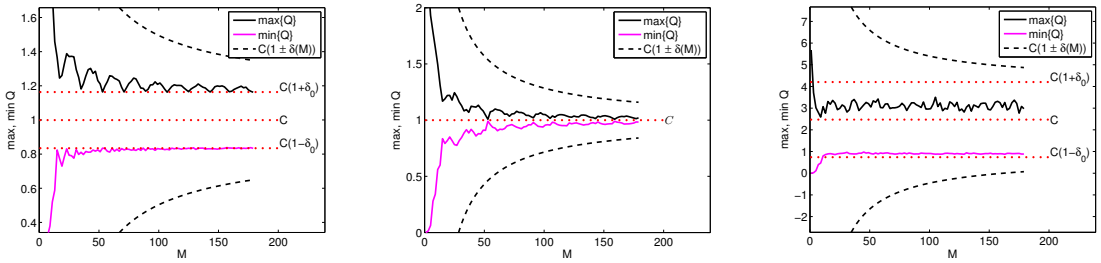


Figure 15: Simulations exploring the asymptotic bounds on the conditioning of the delay-coordinate map. Plotted are the largest and smallest value of $Q(x, y)$ (depicted by $\max\{Q\}$ and $\min\{Q\}$ respectively) attained by the 1000 pairs of x, y for each M. The dotted (red) lines represent the values of $C(1 \pm \delta_0)$ and C, and the dashed (black) lines are the theoretical values of $C(1 \pm \delta(M))$. (a) In this simulation, $A_1 = A_2$ but $\kappa_1 \neq \kappa_2$, thus a plateau on the conditioning is seen. (b) In this simulation, $A_1 = A_2$ and $\kappa_1 = \kappa_2$. As expected, the conditioning number asymptotically reaches 0 as M grows. (c) In this simulation, $A_1 \neq A_2$ and $\kappa_1 \neq \kappa_2$ and the predicted asymptotic values of the conditioning are not tight.

$\max\{Q\}$ and $\min\{Q\}$ as expected from the theorem, and the characteristic shape of these curves seems to qualitatively reflect the empirically observed convergence of the conditioning number.

As confirmation, we also verify the implication of Theorem 5.2.2 that system and measurement combinations can be constructed where the conditioning can be made arbitrarily good with more measurements (akin to the more typical CS results). To show this, we create another system with the same $\mathcal{A}$-eigenvalues and $\mathcal{A}$-eigenvectors as in the previous simulation, with the latter implying that $A_1 = A_2$. For the observation function, we first define $c = V[1, \cdots, 1]^T$, and for each M we let $h = h(M) = \sqrt{\frac{2d}{M}} \frac{c}{\|c\|_2}$ as before. One can verify this choice results in $|v_i^H h|/\|h\|_2 = 1$

for all i, and thus $\kappa_1 = \kappa_2$. The parameters of this experiment are summarized in Table 3.

Index i	1	2	3	4	5	6
θ_i (rad)	2.3129	0.1765	1.4861	—	—	—
$\lvert v_i^H h\rvert^2/\lVert h\rVert_2^2$	1	1	1	—	—	—
$\lambda_i(V^H V)$	1	1	1	1	1	1

Table 3: Parameters for the simulation shown in Figure 15(b). The experiment was chosen such that $A_1 = A_2 = 1$ and $\kappa_1 = \kappa_2 = 1$, so that $\delta_0 = 0$. As the $\mathcal{A}$-eigenvalues are the same as in the previous experiment, ν remains at 5.6954.

With this choice of parameters such that $A_1 = A_2$ and $\kappa_1 = \kappa_2$, Theorem 5.2.2 indicates that $\delta_0 = 0$ so that $\lim_{M\to\infty} \delta(M) = 0$. Figure 15(b) shows the results of running the simulation in the same manner as before. The values of $\max\{Q\}$ and $\min\{Q\}$ clearly converge to C as expected, showing that in this case the conditioning of the embedding can indeed be made arbitrarily good by taking more measurements.

Although Theorem 5.2.2 indicates that a finite limit on the conditioning number is always reached when either $A_1 \neq A_2$ or $\kappa_1 \neq \kappa_2$, this bound is not always tight and the predicted plateau level of $C(1 \pm \delta_0)$ may be conservative. To show this, we construct a similar simulation as above, now setting the $\mathcal{A}$-eigenvectors to be $v_i = \frac{1}{\sqrt{\|a_i\|_2^2+\|b_i\|_2^2}}(a_i + jb_i)$, where $\{a_i, b_i\}$ are randomly constructed vectors in $\mathbb{R}^N$ whose entries are i.i.d. zero-mean Gaussian random variables with a variance of 1. We keep the $\mathcal{A}$-eigenvalues the same and generate h in the same manner as the first simulation shown in Figure 15(a). The specific parameters for this simulation are shown in Table 4, where we see that indeed $A_1 \neq A_2$ and $\kappa_1 \neq \kappa_2$. Figure 15(c) shows the results of running the simulation in the same manner as before. We see that although a limit on the conditioning number is reached as predicted by Theorem 5.2.2, the predicted plateau level of $C(1\pm\delta_0)$ is not tight and the conditioning can be better than that predicted by δ_0.

Index i	1	2	3	4	5	6
θ_i (rad)	2.3129	0.1765	1.4861	—	—	—
$\lvert v_i^H h\rvert^2/\lVert h\rVert_2^2$	1.8138	1.2064	1.1318	—	—	—
$\lambda_i(V^H V)$	1.5316	1.3058	1.1294	0.8372	0.7644	0.4315

Table 4: Parameters for the simulation shown in Figure 15(c). We see that $A_1 = 0.4315$, $A_2 = 1.5316$, $\kappa_1 = 1.1318$ and $\kappa_2 = 1.8138$. Since the $\mathcal{A}$-eigenvalues are the same as in the first simulation shown in Figure 15(a), ν remains the same at 5.6954. We also calculate $\delta_0 = 0.7010$.

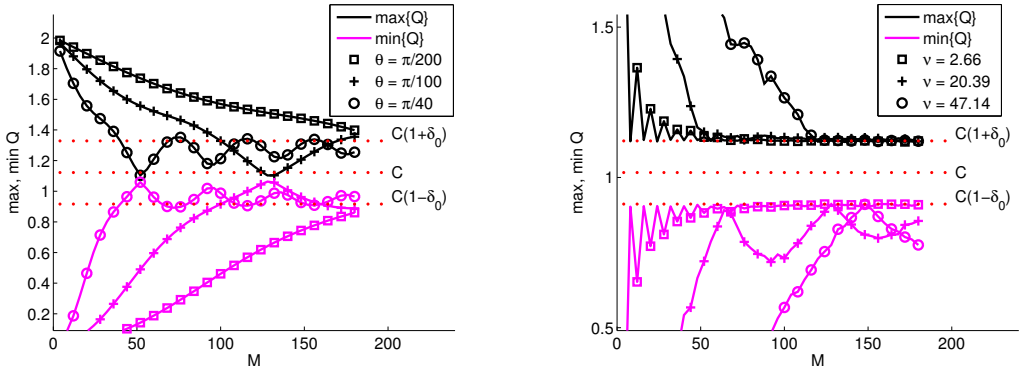


Figure 16: Examining the effect of the $\mathcal{A}$-eigenvalues on the convergence speed of the conditioning. (a) In this simulation, $d = 1$ and we test $\theta = \frac{\pi}{200}, \frac{\pi}{100}$ and $\frac{\pi}{40}$. As expected, the closer θ is to $\pi/2$, the faster the rate of convergence of $\delta(M)$ to δ_0. (b) In this simulation, $d = 3$ and we vary between 3 sets of $\mathcal{A}$-eigenvalues with different values of ν. As expected, the set of eigenvalues that gives the smallest ν provides the fastest rate of convergence of $\delta(M)$ to δ_0 and vice versa.

5.2.3.2 Convergence Speed

In the simulations of the previous section we concentrated on the conditioning limits predicted by Theorem 5.2.2, ignoring issues of the speed of convergence to those limits. Examining the formula for $\delta_1(M)$ in Theorem 5.2.2, we see that the $\mathcal{A}$-eigenvalues (via the parameter ν) affect the convergence speed of $\delta(M)$ to its asymptotic value of δ_0. In particular, the convergence speed scales with $1/\nu$, which is also demonstrated in (33) where the number of measurements $\widehat{M}(\epsilon)$ necessary to get the conditioning δ within ϵ of the best possible value (δ_0) is proportional to ν.

For ease of analysis, we first consider the case where $d = 1$, meaning that $\nu =$

$|\sin(\theta)|^{-1}$ (since $T_s = 1$), where $\pm j\theta$ are the sole $\mathcal{A}$-eigenvalues. In this case, $|\sin(\theta)|^{-1} \geq 1$ with the minimum attained when $\theta = \frac{\pi}{2} + k\pi$ for any integer k. The closer θ is to $\frac{\pi}{2}+k\pi$, the faster the convergence of $\delta(M)$ to δ_0. This is illustrated by the following simulation where the $\mathcal{A}$-eigenvectors are chosen such that $A_1 = A_2$, and the observation function is chosen randomly as in the experiment shown in Figure 15(a) (except with $d = 1$). Figure 16(a) plots $\max\{Q\}$ and $\min\{Q\}$ for $\theta = \frac{\pi}{200}, \frac{\pi}{100}$ and $\frac{\pi}{40}$, showing that Theorem 5.2.2 correctly captures that the convergence speed to the asymptotic value of $C(1 \pm \delta_0)$ varies inversely with the value of θ.

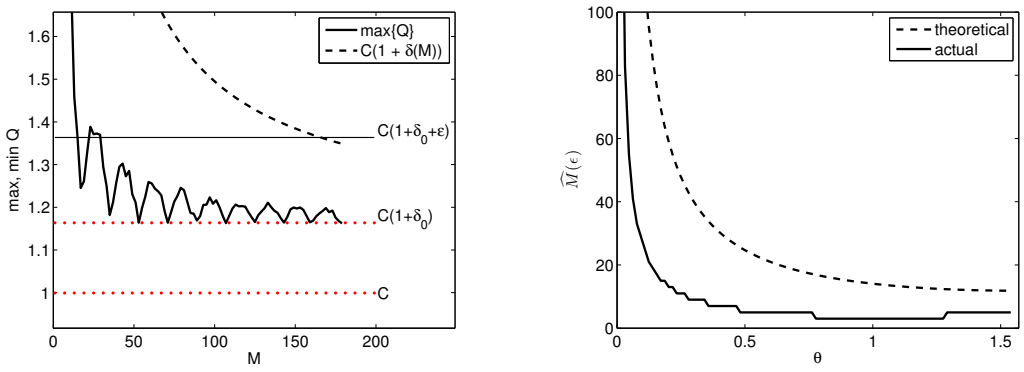
When $d > 1$, the joint relationship of the $\mathcal{A}$-eigenvalues (not just their individual values) determines ν, and subsequently the convergence speed. One can see intuitively in the definition of ν that $\mathcal{A}$-eigenvalues which are maximally spread out should produce favorable convergence speeds. To illustrate this, we generate a simulated system with $d = 3$, choosing the $\mathcal{A}$-eigenvectors such that $A_1 = A_2$, and generating an observation function h randomly (as in the experiment in Figure 15(a)). We also choose three sets of $\mathcal{A}$-eigenvalues: two uniformly random sets, and one set that are slight perturbations of equally spaced points around the unit circle according to $\theta_p = \frac{p\pi}{d+1}$ (the choices of θ_p and their respective ν are given in Table 5).[11] Figure 16(b) shows the results of the simulation, with the $\max\{Q\}$ and $\min\{Q\}$ curves showing clearly that ν indeed controls the speed of convergence of $\delta(M)$ as predicted.

	θ_1	θ_2	θ_3	ν
Set 1 (nearly equal spacing)	0.7836	1.5864	2.3566	2.6619
Set 2 (random)	0.0491	1.5737	2.3490	20.3851
Set 3 (random)	0.0212	1.5684	2.3549	47.1388

Table 5: Choice of $\{\theta_i\}$ (in radians) for the experiment in Figure 16(b) and their respective ν value.

Given that Theorem 5.2.2 seems to be correctly capturing the convergence speed

[11] The slight perturbation is used for plotting convenience so all three curves converge to the same asymptotic value. If exactly equally spaced eigenvalues are used, the attractor is sampled uniformly and the convergent value will be inside $C(1 \pm \delta_0)$, making comparative plots difficult.



(a) (b)

Figure 17: Examining the predicted number of measurements necessary to reach a specified conditioning level. (a) Plotted is the upper-half of Figure 15(a), also indicating $C(1+\delta_0+\epsilon)$ with $\epsilon = 0.2$. (b) In this simulation, we explore how $\widehat{M}(\epsilon)$ (for a fixed $\epsilon = 0.1$) varies with the $\mathcal{A}$-eigenvalues for the system defined in Figure 16(a). We plot the theoretical values of $\widehat{M}(\epsilon)$ (given in (33)) for θ varying from 0 to $\pi/2$ together with its actual values (as described in the text) obtained by running experiments for each θ.

dependence on ν, the last facet of the problem to explore is the tightness of this bound. Specifically, given a system of class $\mathcal{A}(d)$ and an observation function h, it is often of interest to estimate the minimum number of measurements $\left(\widehat{M}(\epsilon)\right)$ needed to ensure that for any $M' \geq M$ the conditioning number $\delta(M')$ is at most ϵ above the asymptotic level of δ_0 (such an estimate is given in (33)). To examine this, we refer back to the simulation shown in Figure 15(a) with parameters given in Table 2. Fixing $\epsilon = 0.2$, Figure 17(a) re-plots $\max\{Q\}$ together with the line $C(1 + \delta_0 + \epsilon)$. Using the given parameters and (33) we calculate that $\widehat{M}(\epsilon) \approx 166$. Note that this value is also the intersection of the curve $C(1 + \delta(M))$ with the line $C(1 + \delta_0 + \epsilon)$. Figure 17(a) shows that $\max\{Q\}$ actually met this tolerance with only around 30 measurements. Thus, although the theoretical value of $\widehat{M}(\epsilon)$ given by (33) is correct, it is pessimistic in at least this particular case.

To demonstrate that the linear dependence of $\widehat{M}(\epsilon)$ on ν is correctly captured in the theorem, we restrict ourselves to $d = 1$. Recall that when $d = 1$, $\nu = |\sin(\theta)|^{-1}$ (since $T_s = 1$) where $\pm j\theta$ are the sole $\mathcal{A}$-eigenvalues. We repeat the simulation shown

in Figure 16(a), this time using 100 values of θ equally spaced between $(0, \pi/2)$. Fixing $\epsilon = 0.1$, for each value of θ we note the value of M where for all $M' > M$, $\max\left\{\frac{\max\{Q\}}{C} - 1,\ 1 - \frac{\min\{Q\}}{C}\right\} < \delta_0 + \epsilon$. We call this value the "actual" $\widehat{M}(\epsilon)$, in contrast to the "theoretical" $\widehat{M}(\epsilon)$ given by (33). Figure 17(b) shows these actual and theoretical values of $\widehat{M}(\epsilon)$ as a function of θ. This comparison shows that while the theoretical $\widehat{M}(\epsilon)$ captures the same trend as the actual $\widehat{M}(\epsilon)$, the theoretical estimate can be pessimistic compared to the empirical values (though it is not clear if the theoretical bounds are achieved by some systems).

5.2.3.3 Stable Embeddings for Dimension Estimation

To demonstrate the value of stable Takens' embeddings, this section will explore a simulated task estimating the dimensionality of an attractor. The *correlation dimension* is a measure of attractor dimension often applied to strange attractors of chaotic systems [68], which corresponds to the actual geometric dimension of regular objects such as the circles and ellipses seen in linear system attractors [81]. To be precise, we first define the *correlation sum* of tolerance ϵ for a set of points $\{x_k\}$ lying on a subset $\mathcal{M}$ and temporally related via the flow (i.e., $x_k = \psi^k x_0$) as

$$C(\epsilon, K) := \frac{2}{K(K-1)} \sum_{p=1}^{K} \sum_{q=p+1}^{K} \Theta(\epsilon - \|F(x_p) - F(x_q)\|_2), \tag{34}$$

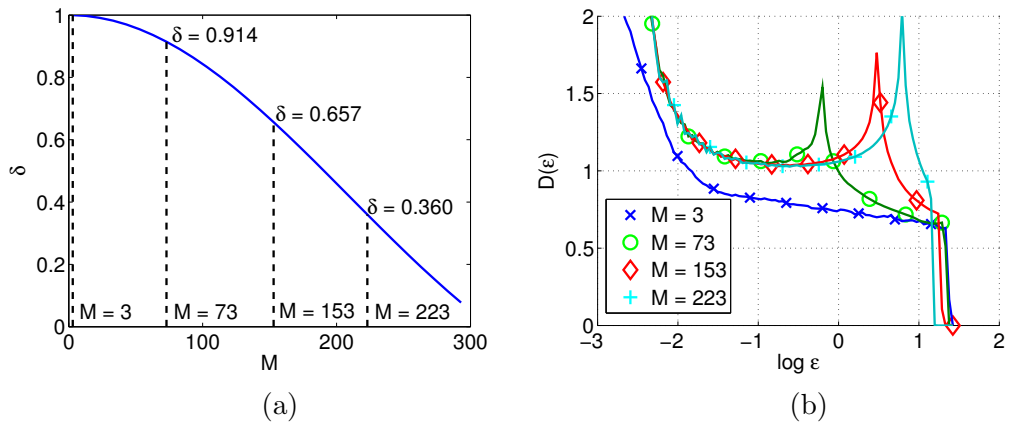
where F is the delay-coordinate map and $\Theta(\cdot)$ is the *Heaviside step function* defined as $\Theta(x) = 0$ if $x \leq 0$ and $\Theta(x) = 1$ if $x > 0$. The *correlation dimension* is defined as $D = \lim_{\epsilon \to 0} \lim_{K \to \infty} \frac{\partial \log C(\epsilon, K)}{\partial \log \epsilon}$. This makes intuitive sense as in the limit of small ϵ and large K, we expect $C(\epsilon, K)$ to scale like $C(\epsilon, K) \propto \epsilon^{-D}$, where D is the dimension of the subset $\mathcal{M}$ in question. Theoretically, one way to estimate correlation dimension is to plot the graph of $\log C(\epsilon, K)$ against $\log \epsilon$ for a large value of K, then simply read off the gradient for small values of $\log \epsilon$. In the absence of noise and with a topology preserving Takens' embedding (i.e. $M > 2d$), this estimate should be as good as if one had access to the hidden system state. However, when noise is present,

small values of $\log \epsilon$ will be capturing the noise characteristics and overestimating the attractor dimension. A common approach in this case is to plot the local gradient $D(\epsilon) = \frac{\partial \log C(\epsilon,K)}{\partial \log \epsilon}$ against $\log(\epsilon)$ for a large value of K and read off an estimate of the correlation dimension D from a plateau in the graph, preferably in the regime of small ϵ.

In this section, we use the above approach to estimate the correlation dimension of linear system attractors $\mathcal{M}$ in the reconstruction space $\mathbb{R}^M$. For this simulation construct a linear dynamical system of class $\mathcal{A}(1)$ with $N = 100$, $\mathcal{A}$-eigenvalue $\theta = \frac{\pi}{300}$ and $\mathcal{A}$-eigenvector $v = [1,\ j]^T$ (resulting in $A_1 = A_2$ and a circular attractor). We also choose $h = [1,\ 1]^T$, implying that $\kappa_1 = \kappa_2$ and subsequently $\delta_0 = 0$. Figure 18(a) shows that the actual conditioning[12] of F approaches zero as we increase M. To simulate noisy measurements, we corrupt the resulting time series formed by h by adding white gaussian noise with zero mean and standard deviation $\sigma = 0.05$ (to give an SNR of about $32dB$).

Figure 18(b) shows the plots of $D(\epsilon)$ against $\log(\epsilon)$ with a number of delays $M = 3, 73, 153, 223$. For the graph corresponding to $M = 223$, a plateau is easily seen between $-1 < \log \epsilon < 0$, and corresponding to a correct dimension estimate of approximately 1. We observe that by taking more measurements (i.e., improving the conditioning of the embedding), the estimate of the correlation dimension also improves. Moreover, the width of the plateau region where we read off the correlation dimension estimate increases with increasing M, thus making its estimate more precise. Note that when we take the minimum number of measurements $M = 3$ required by Takens' Theorem, there is no discernible plateau region in Figure 18(b) for us to estimate the correlation dimension, and even the most reasonable estimate near $\log \epsilon = 1$ is less accurate than with the estimates produced by the embeddings with

[12]By actual conditioning, we mean the empirical value $\delta = \max\left\{\frac{\max\{Q\}}{C} - 1, 1 - \frac{\min\{Q\}}{C}\right\}$, for Q defined in Section 5.2.3.



(a) (b)

Figure 18: Estimating the correlation dimension of a circular attractor $\mathcal{M}$ of a linear system of class $\mathcal{A}(1)$. (a) The conditioning of the stable embedding decreases with increasing number of measurements M. (b) The graphs of $D(\epsilon)$ for the various M considered are plotted against $\log \epsilon$. The correlation dimension estimate can be read off the plateaus in these graphs. These plateau regions become more distinct with increasing M (improving conditioning), and appear to converge to a value near the true dimension of 1.

better conditioning.

5.2.4 Conclusion

We established that a delay-coordinate map (using linear observation functions) can form a stable embedding for all pairs of points on the attractor of a linear dynamical system of class $\mathcal{A}(d)$. The explicit, deterministic and non-asymptotic sufficient conditions we give for this stable embedding yield several observations about the embedding itself and favorable properties of system and measurement pairs. For example, for many system and measurement pairs, the conditioning number $\delta(M)$ reaches a non-zero asymptotic value of δ_0 with increasing M. This "plateau effect" is in contrast with typical CS results where the conditioning of the stable embedding can be continually improved by increasing the number of measurements. Furthermore, the convergence speed of the embedding quality to this limit is governed by the joint relationship of the system eigenvalues, which capture the relative speed with which the system explores the different dimensions of the state space (i.e., more diversity in these speeds results in faster convergence). Finally, we also see that the minimum

number of delays M of the delay-coordinate map scales linearly with the attractor dimension but is independent of the system dimension. This is again in contrast with typical CS results, where the number of compressive measurements also scales logarithmically with the system dimension (but parallels recent improvements in these bounds for the stable embedding of manifolds [39]).

While the comparisons with standard CS results reveal these interesting and non-intuitive technical differences between the results in each case, these discrepancies actually point to a much deeper difference in the problem setups that must be appreciated when embedding attractors of dynamical systems. Perhaps the easiest way to see this is to consider that in the present case of delay-coordinate maps, while the number of measurements doesn't scale with the ambient system dimension, the total number of measurements may in fact have to be larger than the system dimension ($M > N$) in order to make a particular conditioning guarantee. In the typical CS case, this would of course be a ridiculous proposition. If the RIP property required ($M > N$) random measurements (e.g., due to very large constants in the typical sufficient conditions), one would likely abandon the CS strategy and simply take N uncoded measurements (e.g., in the canonical basis). However, in the case of delay-coordinate maps for dynamical systems, this luxury is simply not available. For example, the observers often do not have any control over the choice of observation function h, and in these cases cannot simply change the way the system is measured. But, more importantly, even if we were given complete control over h, it is only a "seed" that is used in producing the whole measurement process. One can view the entire set of measurements as being generated by repeatedly forcing this observation function through the dynamics of the system (seen explicitly in writing the delay-coordinate map in (30)). Said another way, because there is only a single observation function for the system, the total measurement process for a delay-coordinate map is beholden to the dynamics of the system itself to provide sufficient diversity

to make the measurements informative. Therefore, even with complete control over the observation function, delay-coordinate maps represent a highly restricted total measurement process that cannot be completely controlled (without access to and control over the system that is hidden and in need of measurement).

Characterizing the delay-coordinate map embeddings for attractors of linear dynamical systems with linear observation functions is a subset of the more general problem of characterizing these embeddings for attractors of nonlinear systems and general observation functions. From the results here, we conclude that there is reason to be optimistic that similar stability results can be obtained for this more general case of interest. Furthermore, these results also lead us to conclude that there are several issues that differ from standard CS results and will need to be carefully considered in any generalization.

5.3 Nonlinear Dynamical Systems and Nonlinear Observation Functions

In this section, we move from characterizing delay-coordinate embeddings for attractors of linear dynamical systems with linear observation functions to the more general problem of characterizing these embeddings for attractors of nonlinear systems and general observation functions. As our results depend on the types of measurement functions used, we define a *class* of measurement functions as follows. First, consider a set of basis measurement functions $h_p : \mathbb{R}^N \to \mathbb{R}$ for $p \in [P]$ that span a P-dimensional subspace of the (infinite-dimensional) space of all smooth functions from $\mathbb{R}^N$ to $\mathbb{R}$. This set of functions $\{h_p\}$ forms the basis for the class of measurement functions that we are considering. For example, if we are considering the class of linear functions, then $P = N$ (where N is the ambient dimension) and $h_p = e_p^T$ where e_p is the p-th canonical vector. If we are considering the class of polynomials up to a certain degree k, then the set of basis functions can be the set of monomials up to that degree k.

An element chosen from this class of measurement function is denoted by h_α for some $\alpha \in \mathbb{R}^P$ defined as $h_\alpha := \sum_{p=1}^{P} \alpha_p h_p$. For $\alpha \in \mathbb{R}^P$, we define the delay-coordinate map with measurement function h_α (and with M delays) as

$$F_\alpha(x) := \begin{bmatrix} h_\alpha(x) \\ h_\alpha\left(\psi^{-1}(x)\right) \\ \vdots \\ h_\alpha\left(\psi^{-M+1}(x)\right) \end{bmatrix} = \sum_{p=1}^{P} \alpha_p F_p(x),$$

where F_p is defined as the delay-coordinate vector formed with the p-th basis function

$$F_p(x) := \begin{bmatrix} h_p(x) \\ h_p\left(\psi^{-1}(x)\right) \\ \vdots \\ h_p\left(\psi^{-M+1}(x)\right) \end{bmatrix}.$$

We first present our stable embedding result for the relatively easier case of linear measurement functions in Section 5.3.1. Then, we will move on to the more general case of nonlinear measurement functions in Section 5.3.2.

5.3.1 Linear Measurement Functions

As previously discussed, the basis functions for the class of linear measurement functions are $h_p = e_p^T$ for $p \in [N]$ (so $h_\alpha = \alpha^T$). We assume that we have a dynamical system given by the (smooth) flow $\psi = \psi_{T_s}$ with sampling time T_s and whose states are lying on a submanifold $\mathcal{M}$ of $\mathbb{R}^N$.

Before introducing our stable embedding result, we require a few additional definitions. For any state $x \in \mathcal{M}$, define the $M \times N$ matrix G_x as

$$G_x = (F_1(x) \mid \cdots \mid F_N(x)) = \begin{pmatrix} x^T \\ \vdots \\ \psi^{-M+1}(x)^T \end{pmatrix}.$$

Next, define the trajectory vector $\widetilde{g}(x) \in \mathbb{R}^{MN}$ starting at point $x \in \mathcal{M}$ as

$$\widetilde{g}(x) := \left[x^T, \psi^{-1}(x)^T, \cdots, \psi^{-M+1}(x)^T\right]^T.$$

Observe that G_x is basically the matrix version of $\widetilde{g}(x)$ and that $F_\alpha(x) = \sum_{i=1}^{N} \alpha_i F_i(x) = G_x\alpha$. We also define $\widetilde{g}(\mathcal{M}) \subset \mathbb{R}^{MN}$ as the set of all trajectories in $\mathcal{M}$, i.e.,

$$\widetilde{g}(\mathcal{M}) := \{\widetilde{g}(x) \,:\, x \in \mathcal{M}\}.$$

Notice that since the first N entries of $\widetilde{g}(x)$ are distinct for all $\widetilde{g}(x) \in \widetilde{g}(\mathcal{M})$ and ψ is a smooth function, the trajectory manifold $\widetilde{g}(\mathcal{M})$ is diffeomorphic to the ambient space manifold $\mathcal{M}$.

For any $G \in \mathbb{R}^{M\times N}$, define $r(G) := \frac{\|G\|_F^2}{\|G\|_2^2}$ the *soft-rank* of the matrix G. The soft rank is a refined metric for the rank, and by observing that the soft-rank is the ratio of the ℓ_2 and ℓ_∞ norms of the vector of singular values, we have

$$1 \le r(G) \le \operatorname{rank}(G).$$

The upper bound is attained when all the singular values of G are equal, while the lower bound is attained when only one of the singular values is significant while the rest are zero.

In the stable embedding results that follow, we shall see that the soft-rank of the matrices $G_x - G_y$ for all pairs $x, y \in \mathcal{M}$ controls the number of delays necessary to ensure stable embedding of $\mathcal{M}$. Under the assumptions that there are no periodic orbits of period less than M on $\mathcal{M}$ and $M \le N$, it is easy to see that $\operatorname{rank}(G_x - G_y) = M$ for all pairs of $x, y \in \mathcal{M}$. Under these assumptions , $r(G_x - G_y) = cM$ for some $c \in (\frac{1}{M},\ 1]$. We shall see that the required number of measurements M scales *inversely* with c, therefore we will want c to be as close to 1 as possible. We remark that via linear algebra, $c \approx 1$ whenever the rows of $G_x - G_y$ (i.e., the chords $x - y, \psi^{-1}(x) - \psi^{-1}(y), \cdots$) have approximately equal length and are approximately orthogonal.

With all the required notations now defined, we finally present our stable embedding result for linear measurement functions.

Theorem 5.3.1. *Suppose we have a dynamical system described by a discrete flow ψ with sampling time T_s and whose states lie on a $D_{\mathcal{M}}$-dimensional compact submanifold $\mathcal{M}$ of $\mathbb{R}^N$. Also, suppose that the trajectory manifold $\widetilde{g}(\mathcal{M})$ has volume $V_{\widetilde{g}(\mathcal{M})}$, geodesic covering regularity $R_{\widetilde{g}(\mathcal{M})}$ and second fundamental form bounded by $\frac{1}{\tau_{\widetilde{g}(\mathcal{M})}}$.*[13] *Let $\alpha \in \mathbb{R}^N$ be an i.i.d. Rademacher sequence, and let ρ and δ be a predetermined failure probability and stable embedding conditioning respectively. Assume that*

$$\inf_{x \neq y \in \mathcal{M}} r(G_x - G_y) \quad \gtrsim \quad \frac{1}{\delta^2}\left[D_{\mathcal{M}} \log\left(\frac{R_{\widetilde{g}(\mathcal{M})}\sqrt{D_{\mathcal{M}}}PM}{\delta \tau_{\widetilde{g}(\mathcal{M})}}\right) + \log\left(\frac{V^2_{\widetilde{g}(\mathcal{M})}}{\rho}\right)\right]. \quad (35)$$

Then, with probability exceeding $1-\rho$, the delay-coordinate map with M delays F_α is a stable embedding of $\widetilde{g}(\mathcal{M})$, i.e., for all $x, y \in \mathcal{M}$

$$(1-\delta) \leq \frac{\|F_\alpha(x) - F_\alpha(y)\|_2^2}{\|\widetilde{g}(x) - \widetilde{g}(y)\|_2^2} \leq (1+\delta). \quad (36)$$

Proof. See Appendix E.2. □

Our result describes how the soft-rank of $G_x - G_y$ in (35), which captures the geometry of the manifold and the dynamics of the system over M time steps, dictates the stable embedding conditioning of the delay-coordinate map. Thus, this result is different from the typical Takens' embedding result in two distinct ways. First, while only the *topology* of the manifold (basically the fact that the system states lie on a manifold and the dimension of the manifold) is utilized in the typical Takens' embedding result, the *geometry* of the manifold (basically the trajectory manifold characteristics $V_{\mathcal{M}}, R_{\widetilde{g}(\mathcal{M})}, \tau_{\widetilde{g}(\mathcal{M})}$) is an important consideration for ensuring stable embeddings here. Second, the typical Takens' result does not say what happens when we have more measurements beyond the minimum required for an embedding. Here, an increased number of measurements will change the soft-rank of $G_x - G_y$ and thus the stable embedding conditioning may improve in certain scenarios (that will be described below).

[13] *By using the joint-manifold analysis in [45], we can relate the trajectory manifold characterizations to the same characterizations of the ambient space manifold.*

The atypical condition (35) demands further elaboration. First, suppose the *infimum soft-rank* term $\inf_{x\neq y\in\mathcal{M}} r(G_x - G_y) = M$. Then, (35) yields a typical stable manifold embedding condition whereby stable embedding is achieved whenever M scales with the dimension of the manifold and logarithmically with other characterizations of the manifold. Moreover, the conditioning δ of the embedding improves with increasing M.[14]

However, the infimum soft-rank equals M only when the rows of the matrices $G_x - G_y$ (i.e., the chords $x - y, \psi^{-1}(x) - \psi^{-1}(y), \cdots$) have the same length and are mutually orthogonal for all $x, y \in \mathcal{M}$. As expected, these length-preserving and orthogonality requirements can be very stringent, and any deviation from these requirements will result in a less favorable scaling of the infimum soft-rank with M (i.e., we get $\inf_{x\neq y\in\mathcal{M}} r(G_x - G_y) = cM$ for $c < 1$). The deviation from these requirements are related to the irrelevancy and redundancy conditions described in Section 5.1.2. First, irrelevancy occurs for a dynamical system whose inverse flow ψ^{-1} has a large maximal Lyapunov exponent. In this case, the requirement that the chords retain almost the same length will be easily violated for any decent sampling time T_s and for x, y close to one another.[15] Second, redundancy occurs when inverse flow ψ^{-1} has a small maximal Lyapunov exponent or T_s is small. Then, the M consecutive chords forming the rows of $G_x - G_y$ can more or less be approximated by vectors on the tangent plane of $\mathcal{M} - \mathcal{M}$ at $x - y$. Since $\dim(\mathcal{M} - \mathcal{M}) \leq 2D_{\mathcal{M}}$ which could potentially be much less than M, $r(G_x - G_y) = cM$ with $c \ll 1$. We will need many more measurements M to "break out" of the tangent plane to achieve $r(G_x - G_y) = cM$ with any decent c (i.e. $c \approx 1$). It is useful to add that if the

[14] We remark that the improvement of δ with M matches what was observed by Casdagli et al. [31] that with more measurements, the distortion (see Section 5.1.2 for a short description of this term) by the delay-coordinate map is decreased.

[15] When the inverse flow ψ^{-1} has Lyapunov exponent λ, we have $\|\psi^{-m}(x) - \psi^{-m}(y)\|_2 \approx e^{m\lambda}\|x - y\|_2$ for x, y close enough. Therefore, the chords going down the matrix $G_x - G_y$ grows exponentially in length, severely violating the length-preserving requirement.

manifold $\mathcal{M}$ indeed only lie on a low-dimensional subspace of $\mathbb{R}^N$, then the infimum soft-rank will be *upper bounded* by the dimension of this subspace and will *not* scale with M. This remark is also true when M exceeds the ambient dimension N, in which case the infimum soft-rank will plateau at N. Notice that this imposes the limitation that $M \leq N$ for this result to be useful, which in turn restricts the class of dynamical systems that we can consider. Notably, this result is not adequate for dynamical systems residing on low-dimensional ambient spaces (e.g., the Lorenz system). Nonetheless, we shall see in the next section that we can break this ambient space dimension "barrier" by considering nonlinear measurement functions.

Another important observation is that instead of embedding state-space vectors $x \in \mathcal{M}$, we are in fact embedding trajectory vectors $\widetilde{g}(x) \in \widetilde{g}(\mathcal{M}) \subset \mathbb{R}^{MN}$. Also, observe that the variables on the right-hand side of (35) are also based on the geometry of $\widetilde{g}(\mathcal{M})$ (i.e., the terms $R_{\widetilde{g}(\mathcal{M})}, V_{\widetilde{g}(\mathcal{M})}$ and $\tau_{\widetilde{g}(\mathcal{M})}$). Maintaining distances between trajectories of the dynamical system in the reconstruction space may be advantageous for some time-series applications. Recall the simple time-series prediction algorithm described in Section 5.1.1. The first step in the algorithm is to form a neighborhood of radius ϵ, denoted by $\mathcal{U}_\epsilon(F(x(t)))$, around the delay-coordinate vector of the current time-series value $F(x(t))$. If F is a stable embedding of the trajectory manifold, then the points $F(x(t_k)) \in \mathcal{U}_\epsilon(F(x(t)))$ correspond to trajectories $\widetilde{g}(x(t_k))$ that are close to the reference trajectory $\widetilde{g}(x(t))$. As we use the average of $s(t_k + T_s)$ as the predicted time-series value, ensuring that trajectories are close is stronger than simply ensuring that the states $x(t_k)$ are close to the reference state $x(t)$ (as would be ensured by a typical Takens' embedding).

However, in other applications, we may want a stable embedding of the ambient space manifold $\mathcal{M}$ instead. Since there there is a diffeomorphism (hence isomorphism) between the ambient and the trajectory manifold, our result can be converted into a stable embedding result for the ambient space manifold. Understandably, this

translation comes with a degradation of the stable embedding conditioning δ. To see this, we begin by writing

$$\|\widetilde{g}\,(x) - \widetilde{g}\,(y)\,\|_2^2 = \sum_{m=1}^{M} \|\psi^{-m+1}(x) - \psi^{-m+1}(y)\|_2^2,$$

for some $x \neq y \in \mathcal{M}$. Since the flow ψ is smooth, it is not absurd to assume that the inverse flow ψ^{-1} is bi-Lipschitz, meaning that we can find $C_u \geq C_l > 0$ such that

$$C_l\|x-y\|_2^2 \leq \|\psi^{-1}(x) - \psi^{-1}(y)\|_2^2 \leq C_u\|x-y\|_2^2,$$

so that

$$\sum_{m=0}^{M-1} C_l^m\|x-y\|_2^2 \leq \sum_{m=1}^{M} \|\psi^{-m+1}(x) - \psi^{-m+1}(y)\|_2^2 \leq \sum_{m=0}^{M-1} C_u^m\|x-y\|_2^2.$$

Now, suppose $C_l < 1$ (if not, all trajectories evolve to a fixed point) and $C_u > 1$ (if not, the dynamical system is unbounded which violates the fact that we are lying on a compact manifold). By using the formula for geometric sums, we have for all $x, y \in \mathcal{M}$,

$$C(1-\delta_{\widetilde{g}(\mathcal{M})}) := \frac{C_l^M - 1}{C_l - 1} \leq \frac{\|\widetilde{g}\,(x) - \widetilde{g}\,(y)\,\|_2^2}{\|x-y\|_2^2} \leq \frac{C_u^M - 1}{C_u - 1} =: C(1+\delta_{\widetilde{g}(\mathcal{M})}), \qquad (37)$$

where $C := \frac{1}{2}\left(\frac{C_l^M-1}{C_l-1} + \frac{C_u^M-1}{C_u-1}\right)$ and $\delta_{\widetilde{g}(\mathcal{M})} := \left(\frac{C_u^M-1}{C_u-1} - \frac{C_l^M-1}{C_l-1}\right)\left(\frac{C_u^M-1}{C_u-1} + \frac{C_l^M-1}{C_l-1}\right)^{-1}$. Putting (37) with the trajectory manifold embedding statement (36), we obtain a stable embedding of the ambient space manifold

$$C(1-\delta') \leq C(1-\delta_{\widetilde{g}(\mathcal{M})})(1-\delta) \leq \frac{\|F_\alpha(x) - F_\alpha(y)\|_2^2}{\|x-y\|_2^2} \leq C(1+\delta_{\widetilde{g}(\mathcal{M})})(1+\delta) = C(1+\delta'),$$

but with a worse conditioning $\delta' := (\delta_{\widetilde{g}(\mathcal{M})} + \delta + \delta_{\widetilde{g}(\mathcal{M})}\delta)$.

Notice that by moving the stable embedding from the trajectory manifold to the ambient space manifold, the added conditioning $\delta_{\widetilde{g}(\mathcal{M})}$ imposes a lower bound on the achievable overall stable embedding conditioning. While it is possible to decrease the conditioning δ by increasing M, it is not difficult to see that $\delta_{\widetilde{g}(\mathcal{M})} \to 1$ when $M \to \infty$ (under the assumption that $C_l < 1$ and $C_u > 1$). This means that even by increasing

M (and thus improving the infimum soft-rank), the stable embedding conditioning of $\mathcal{M}$ (not $\widetilde{g}(\mathcal{M})$) by delay-coordinate maps reaches a limit imposed $\delta_{\widetilde{g}(\mathcal{M})}$.

Requirements for the minimum degradation of the stable embedding conditioning caused by mapping trajectory vectors down into ambient space usually run contrary to requirements for the best infimum soft-rank in (35). From its definition, we see that the added conditioning $\delta_{\widetilde{g}(\mathcal{M})}$ is dependent on C_l and C_u, which in turn are dependent on how fast $\|\psi^{-m}(x)-\psi^{-m}(y)\|_2^2$ can grow or shrink with m. When the sampling time and/or the Lyapunov exponent of the inverse flow is small, the trajectory and ambient-space vectors are scalar multiplies of one another. This implies that there is minimum degradation of the stable embedding conditioning going from trajectory to ambient space. However, the soft-rank of $G_x - G_y$ suffers because the rows of the matrix $G_x - G_y$ are very similar to one another which violates the orthogonality condition. On the other hand when either the sampling time or the Lyapunov exponent of the inverse flow are large, the consecutive chords in $G_x - G_y$ rapidly decorrelates. When this occurs, the soft-rank of $G_x - G_y$ may improve due to the diversity but the overall conditioning is degraded when we passed from trajectory space to ambient space.

Lastly, we remark that our results mirror the notion that "almost every measurement function provides an embedding" appearing in a typical Takens' embedding statement (e.g., Theorem 5.1.2, and see also [120]). This notion is represented by the random vector of coefficients $\alpha \in \mathbb{R}^N$. Because F_α is a stable embedding with high probability on the coefficients α, this means that most measurements functions in the space of functions defined by $\{h_\alpha = \alpha^T \mid \alpha \in \{-1, +1\}^N\}$ can result in a stable embedding by delay-coordinate maps. In fact, the class of measurement functions considered can be vastly expanded. This will be discussed briefly in Section 5.3.4.

5.3.2 Nonlinear Measurement Functions

Now that we have a general feel of the stable Takens' Embedding result for linear measurement functions, let us move on to the general case of nonlinear measurement functions. In order to describe this result, we require a few additional notations. Again, we assume that we have a dynamical system given by the flow $\psi = \psi_{T_s}$ with sampling time T_s and whose states are lying on a submanifold $\mathcal{M}$ of $\mathbb{R}^N$. Recall that we are considering measurement functions of the form $h_\alpha = \sum_{p=1}^{P} \alpha_p h_p$ where $h_p : \mathbb{R}^N \to \mathbb{R}$ are basis functions (not necessarily linear) of a particular class of functions of interest. First, let $h : \mathbb{R}^N \to \mathbb{R}^P$ be the vectorized version of the basis functions $\{h_p\}$ so that for any $x \in \mathcal{M}$,

$$h(x) = [h_1(x),\ \cdots,\ h_P(x)]^T .$$

We can view the operation of h on the vectors $x \in \mathcal{M}$ as bringing them from ambient space $\mathbb{R}^N$ to a higher-dimensional *kernel* space $\mathbb{R}^P$, with $P \geq N$. In what follows, we assume that the basis functions $\{h_p\}$ are chosen such that h is a diffeomorphism from $\mathcal{M}$ to $h(\mathcal{M})$.[16] Notice that we have an equivalent flow of the dynamical system in the kernel space given by $\bar{\psi} := h \circ \psi \circ h^{-1}$ such that

$$\begin{aligned} & h(x(t+T_s)) = \bar{\psi}\left(h(x(t))\right) \\ \Leftrightarrow\ & h(x(t+T_s)) = h \circ \psi \circ h^{-1} \circ h(x(t)) \\ \Leftrightarrow\ & x(t+T_s) = \psi(x(t)). \end{aligned}$$

As in the previous section, we define the matrix G_x for any $x \in \mathcal{M}$ as

$$G_x := (F_1(x) \mid \cdots \mid F_N(x)) = \begin{pmatrix} h(x)^T \\ \vdots \\ h\left(\psi^{-M+1}(x)\right)^T \end{pmatrix}.$$

[16] This is not an unreasonable assumption since for a generic choice of $\{h_p\}$, we can invoke Whitney's embedding theorem [120] to say that if $P > 2D_{\mathcal{M}}$ where $D_{\mathcal{M}}$ is the dimension of the system manifold $\mathcal{M}$, then h is an embedding of $\mathcal{M}$.

Instead of the trajectory vector, we now define the *kernelized trajectory vector* $\bar{g}(x) \in \mathbb{R}^{MP}$ as

$$\bar{g}(x) := \left[h(x), h\left(\psi^{-1}(x)\right), \cdots, h\left(\psi^{-M+1}(x)\right)\right]^T.$$

Again, observe that G_x is basically the matrix version of $\bar{g}(x)$ and that $F_\alpha(x) = \sum_{i=1}^{P} \alpha_i F_i(x) = G_x\alpha$. We also define $\bar{g}(\mathcal{M}) \subset \mathbb{R}^{MP}$ as the set of all kernelized trajectories in $\mathcal{M}$:

$$\bar{g}(\mathcal{M}) := \{\bar{g}(x) : x \in \mathcal{M}\}.$$

Notice that since the first P entries of $\bar{g}(x)$ are distinct for all $\bar{g}(x) \in \bar{g}(\mathcal{M})$ (because we chosen h to be diffeomorphic on $\mathcal{M}$) and ψ is a smooth function, $\bar{g}(\mathcal{M})$ is diffeomorphic to $\mathcal{M}$. As we shall see, we will again be concerned with the soft-rank of the matrix $G_x - G_y$.

With all the necessary notations defined, our result for general measurement functions is as follows:

Theorem 5.3.2. *Suppose we have a dynamical system described by a discrete flow ψ with sampling time T_s and whose states lie on a $D_{\mathcal{M}}$-dimensional compact submanifold $\mathcal{M}$ of $\mathbb{R}^N$. Also suppose that the kernelized trajectory manifold $\bar{g}(\mathcal{M})$ has volume $V_{\bar{g}(\mathcal{M})}$, geodesic covering regularity $R_{\bar{g}(\mathcal{M})}$ and second fundamental form bounded by $\frac{1}{\tau_{\bar{g}(\mathcal{M})}}$.*[17] *Let $\alpha \in \mathbb{R}^P$ be an i.i.d. Rademacher sequence and let ρ and δ be a predetermined failure probability and stable embedding conditioning respectively. Assume that*

$$\inf_{x \neq y \in \mathcal{M}} r(G_x - G_y) \gtrsim \frac{1}{\delta^2}\left[D_{\mathcal{M}} \log\left(\frac{R_{\bar{g}(\mathcal{M})}\sqrt{D_{\mathcal{M}}}PM}{\delta\tau_{\bar{g}(\mathcal{M})}}\right) + \log\left(\frac{V^2_{\bar{g}(\mathcal{M})}}{\rho}\right)\right]. \quad (38)$$

Then, with probability exceeding $1-\rho$, the delay-coordinate map with M delays F_α is a stable embedding of $\bar{g}(\mathcal{M})$, i.e., for all $x, y \in \mathcal{M}$

$$(1-\delta) \leq \frac{\|F_\alpha(x) - F_\alpha(y)\|_2^2}{\|\bar{g}(x) - \bar{g}(y)\|_2^2} \leq (1+\delta). \quad (39)$$

[17] *Again by using the joint-manifold analysis in [45], we can relate the kernelized trajectory manifold characterizations to the same characterizations of the ambient space manifold.*

Proof. See Appendix E.2. □

We observe a similar condition on the infimum soft-rank (38) (compared to we have already seen in (35)) except that the matrices $G_x - G_y$ are now defined with the kernel space vectors (instead of trajectory space vectors) and the terms on the right-hand side of (38) depend on characteristics of the kernelized trajectory manifold $\bar{g}(\mathcal{M})$ (instead of the trajectory manifold $\widetilde{g}(\mathcal{M})$). Therefore, much of the discussion of this quantity carries over from the linear measurement functions case. Recall that $r(G_x - G_y) = c\,\mathrm{rank}\,(G_x - G_y)$ with $c \leq 1$. It is shown in [120] that if a monomial basis of degree up to $2M$ is used for the measurement basis functions $\{h_p\}$, then $\mathrm{rank}\,(G_x - G_y) = M$.[18] The favorable case of $c = 1$ occurs when the rows of $G_x - G_y$ are of equal length and are mutually orthogonal. Again, these requirements are typically hard to attain. If the sampling time T_s is too low and/or the Lyapunov exponent of the inverse (kernel space) flow $\bar{\psi}$ is small (i.e., the redundancy condition), then the orthogonality condition will be violated. If the sampling time is large and/or the Lyapunov exponent of the inverse (kernel space) flow $\bar{\psi}$ is large (i.e., the irrelevancy condition), then the length preservation condition will be violated.

A valid question to ask here is whether there are advantages from using nonlinear measurement functions compared to linear ones. First, many naturally occurring measurement functions are nonlinear and the set of linear functions is a subset of the set of nonlinear functions. Thus, this theorem is a more general formulation of the stable Takens' theorem. Another major advantage is that we can remove the restrictive requirement that the number of measurements M must be less than the ambient space dimension N in order to achieve $\mathrm{rank}\,(G_x - G_y) = M$ (although

[18] The attentive reader may realize that using polynomial basis functions will render the condition (38) superfluous. This is because the cardinality of the set of all monomial up to degree $2M$ scales like $P = O(N^M)$, and thus an M term will appear on the right-hand side of (38). Fortunately, as we shall discuss in Section 5.3.4, the logarithmic dependence on P and M on the right-hand side is simply an artifact of our proof technique and can be removed.

for now, we still require that there are no periodic orbits of period less than M in $\mathcal{M}$). Recall that this $M \leq N$ requirement for the linear functions case excluded attractors residing low-dimensional ambient space (e.g., the Lorenz attractor lying in three-dimensional space).

We may hope that by moving the system dynamics from ambient space to a well chosen kernel space, we can improve the infimum soft-rank. However, this is unlikely to occur since the Lyapunov exponent is an invariant characterization of dynamical systems that does not change under changes in coordinates (e.g., moving from ambient to kernel space) [59]. As we have discussed, this means that if the Lyapunov exponent is too small (together with a small sampling time), then even by moving into kernel space, the orthogonality requirement will continue to be violated. Likewise, if the Lyapunov exponent is too large, then the length preservation condition will also continue to be violated. However, the kernel space trajectory manifold may have a more favorable geodesic regularity $R_{\bar{g}(\mathcal{M})}$ and bound on its second fundamental form $\tau_{\bar{g}(\mathcal{M})}$ compared to its ambient space counterpart. This may reduce the quantity on the right-hand side of (38) and may consequently reduce the number of delays M required to achieve stable embedding.

Again, notice that we are showing a stable embedding of the kernel trajectory attractor $\bar{g}(\mathcal{M})$ instead of the ambient space attractor $\mathcal{M}$. Because $\bar{g}(\cdot)$ is a diffeomorphism, we can translate the stable embedding from $\bar{g}(\mathcal{M})$ to $\mathcal{M}$ but suffer a degradation of the conditioning δ in doing so. To see this, we write $\|\bar{g}(x) - \bar{g}(y)\|_2^2 = \sum_{m=0}^{M-1} \|h(\psi^{-m}(x)) - h(\psi^{-m}(y))\|_2^2$. Now, suppose further that h is bi-Lipschitz[19] so that we can write

$$C_{hl}\|\psi^{-m}(x) - \psi^{-m}(y)\|_2^2 \leq \|h(\psi^{-m}(x)) - h(\psi^{-m}(y))\|_2^2 \leq C_{hu}\|\psi^{-m}(x) - \psi^{-m}(y)\|_2^2,$$

[19]This is not far-fetched since h is already assumed a diffeomorphism.

for some constants $C_{hl}, C_{hu} > 0$. Then, we have

$$C_h(1-\delta_h) := C_{hl} \leq \frac{\|\bar{g}(x) - \bar{g}(y)\|_2^2}{\|\widetilde{g}(x) - \widetilde{g}(y)\|_2^2} \leq C_{hu} =: C_h(1+\delta_h),$$

where we used the expansion $\|\widetilde{g}(x) - \widetilde{g}(y)\|_2^2 = \sum_{m=0}^{M-1} \|\psi^{-m}(x) - \psi^{-m}(y)\|_2^2$ and defined $C_h := \frac{1}{2}(C_{hl} + C_{hu})$ and $\delta_h := \frac{C_{hu} - C_{hl}}{C_{hl} + C_{hu}}$. Hence, we obtained an approximate isometry between the kernelized trajectory space and the trajectory space, albeit with a degradation of δ_h depending only on h. From here on, we can apply our previous geometric comparison between the trajectory and ambient space (37) to obtain the following stable embedding statement of the ambient space manifold

$$\begin{aligned} C'(1-\delta') \quad &\leq \quad CC_h(1-\delta_{\widetilde{g}(\mathcal{M})})(1-\delta_h)(1-\delta) \leq \frac{\|F_\alpha(x) - F_\alpha(y)\|_2^2}{\|x-y\|_2^2} \cdots \\ &\cdots \quad \leq CC_h(1+\delta_{\widetilde{g}(\mathcal{M})})(1+\delta_h)(1+\delta) \;=\; C'(1+\delta'), \end{aligned}$$

with $C' := CC_h$ and $\delta' := \delta + \delta_{\widetilde{g}(\mathcal{M})} + \delta_h + \delta\delta_{\widetilde{g}(\mathcal{M})} + \delta\delta_h + \delta_{\widetilde{g}(\mathcal{M})}\delta_h + \delta\delta_{\widetilde{g}(\mathcal{M})}\delta_h$. Recall that $\delta_{\widetilde{g}(\mathcal{M})} \to 1$ as $M \to \infty$ and δ_h does not depend on M. Therefore, as before, we arrive at a quagmire whereby even with increasing M, it is impossible to improve the stable embedding of the state-space manifold $\mathcal{M}$ beyond $\delta_{\widetilde{g}(\mathcal{M})} + \delta_h$.

5.3.3 Stable Embedding Example

It may be useful to validate our theoretical results with a simple example system. Consider the following discrete-time dynamical system with system states at time step n, $x_n \in \mathbb{R}^N$, defined through

$$x_{n+1} = \Phi x_n, \tag{40}$$

where the (linear) flow $\Phi \in \mathbb{R}^{N \times N}$ is a shift matrix, i.e.,

$$\Phi = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & 0 & 1 \\ 1 & 0 & \cdots & \cdots & 0 \end{pmatrix}.$$

Also, suppose the initial condition of the system is given by $x_0 = [1, 0, \cdots, 0]^T$. This system can be thought of as the dynamics of a single point object translating down the entries of a N-dimensional vector. Thus, the system states x_n lie on a one-dimensional manifold $\mathcal{M}$ that is parameterized by the object's location on the vector. We suppose that we only get to observe the time series

$$s_n = \alpha^T x_n,$$

where $\alpha \in \mathbb{R}^N$ is an i.i.d. Rademacher sequence. The goal is to study how well the delay-coordinate vectors formed using the time series s_n is a stable embedding of the trajectory vectors $\widetilde{g}\,(x_n)$ of the dynamical system.

The stable embedding conditioning is dependent on how well the infimum soft-rank scales with the number of measurements M. For $x \neq y \in \mathcal{M}$, we have

$$G_x - G_y = \begin{pmatrix} (x-y)^T \\ (\Phi(x-y))^T \\ \vdots \\ (\Phi^{M-1}(x-y))^T \end{pmatrix},$$

which turns out to be a circulant matrix since Φ is a shift matrix. Moreover, the first row of of $G_x - G_y$ is a vectors of zeros except for a '1' and a '-1'. As shown by the following lemma, the infimum soft-rank of this matrix can be calculated analytically.

Lemma 5.3.1. *For the discrete-time dynamical system described by Φ above, we have* $\inf_{x \neq y \in \mathcal{M}} r(G_x - G_y) \geq \frac{M}{2}$.

Proof. The soft-rank of $G_x - G_y$ is a ratio of the sum of singular values squared over the maximum singular value squared. To calculate its singular values, consider the $M \times M$ symmetric matrix $(G_x - G_y)(G_x - G_y)^T$, where the *eigenvalues* of this matrix correspond to the non-zero singular values squared of the matrix $G_x - G_y$. It is not difficult to deduce that $(G_x - G_y)(G_x - G_y)^T$ is also a circulant matrix, but whose

first row is given by $[2, 0, \cdots, 0, -1, 0 \cdots, 0, -1, 0, \cdots, 0]$. The locations of the '$-1$' on the first row are at the $(|p-q|+1)$-th and $(M-|p-q|+1)$-th position, where p, q correspond to the locations of the '1' and '-1' on the first row of $G_x - G_y$ respectively. We further remark that if the locations of the '-1' coincide (which happens when M is even and $|p-q| = \frac{M}{2}$), we will have a '-2' at the $(|p-q|+1)$-th position instead. Denote by λ_k the k-th eigenvalue of $(G_x - G_y)(G_x - G_y)^H$ (in no particular order). As the eigenvalues of a circulant matrix is simply the (un-normalized) Fourier transform of its first row, we have

$$\lambda_k = 2 - \left(\exp\left(\frac{2\pi}{M}|p-q|k\right) + \exp\left(-\frac{2\pi}{M}|p-q|k\right)\right) = 2 - 2\cos\left(\frac{2\pi}{M}|p-q|k\right).$$

As discussed earlier, the soft-rank of $G_x - G_y$ is given by $r(G_x - G_y) = \frac{\sum_{k=1}^{M}\lambda_k}{\max\{\lambda_k\}}$. First, we have

$$\begin{aligned}\sum_{k=1}^{M}\lambda_k &= 2M - 2\sum_{k=1}^{M}\cos\left(\frac{2\pi}{M}|p-q|k\right) = 2M - 2\,\mathrm{Re}\left\{\sum_{k=1}^{M}\exp\left(\frac{2\pi}{M}|p-q|k\right)\right\} \\ &= 2M - 2\,\mathrm{Re}\left\{e^{\frac{2\pi}{M}|p-q|}\frac{1-e^{2\pi|p-q|}}{1-e^{\frac{2\pi}{M}|p-q|}}\right\} = 2M.\end{aligned}$$

Next,

$$\max_k\{\lambda_k\} = 2 - 2\min_k\left\{\cos\left(\frac{2\pi}{M}|p-q|k\right)\right\} \leq 4.$$

Therefore, we have that for any $x \neq y \in \mathcal{M}$, $r(G_x - G_y) \geq \frac{M}{2}$. □

This lemma tells us that for this system, we get a linear scaling of the infimum soft-rank with M. From (35) in Theorem 5.3.1, this will imply that the stable embedding conditioning δ will scale inversely with the square-root of the number of measurements M.

To test this analysis, we simulated the system described by (40) with $N = 200$. In Figure 19(a), we plotted the minimum soft-rank attained by all pairs of points[20]

[20]Notice that by construction, there are only N distinct states for the system.

taken from the system for each value of M. We see that, indeed, the minimum soft-rank scales linearly with M. Then, we plot in Figure 19(b) the stable embedding conditioning attained by 10 random draws of $h_\alpha = \alpha^T$ for each value of M. The dotted lines are the maximum and minimum conditioning attained while the solid line is the mean. We see that the stable embedding conditioning decreases with increasing M as predicted.

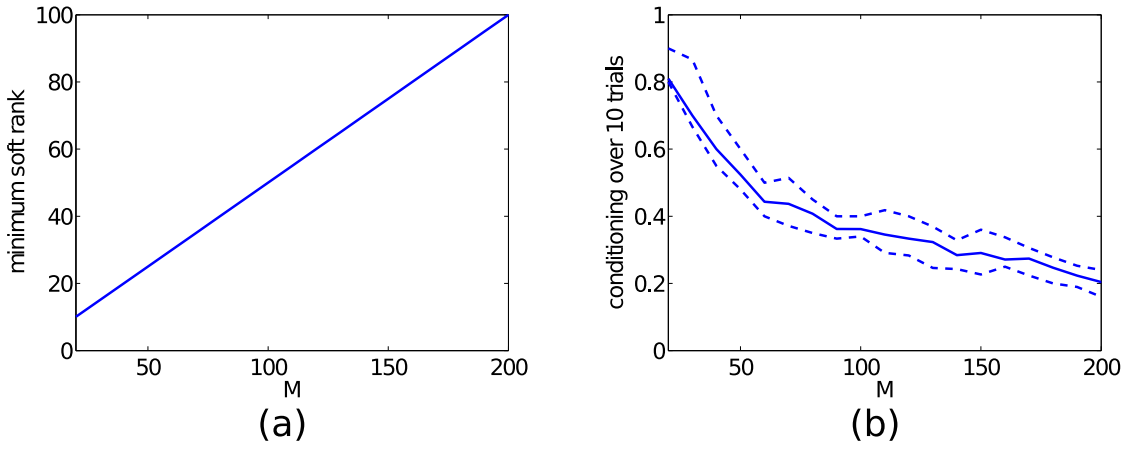


Figure 19: The minimum soft-rank and stable embedding conditioning of the simple one-dimensional dynamical system described in the text. (a) The minimum soft-rank attained by all pairs of points of the system is plotted with increasing M. (b) The stable embedding conditioning attained by 10 random draws of h_α for each M is plotted. The dotted lines are the maximum and minimum conditioning attained while the solid line is the mean.

5.3.4 Discussions

The results we have presented so far can be extended in several useful ways. First, we have considered measurement functions picked uniformly at random from a class of functions given by

$$\left\{ h_\alpha := \sum_p \alpha_p h_p \mid \alpha_p \in \{\pm 1\},\ h_p : \mathbb{R}^N \to \mathbb{R},\ p = 1, \cdots, P \right\},$$

where we recall that $\{h_p\}$ is the set of basis functions that span the P-dimensional measurement function space of interest. However, because α_p can only take values ± 1, the set of possible measurement functions is finite (of size 2^P) and they may not

yield a diverse enough pool of measurement functions to achieve stable embeddings. In light of this short-coming, we can extend our stable Takens' embedding theorems to include a wider class of measurement functions. Specifically, we can consider random measurement functions of the form $h_\alpha = \sum_p \alpha_p h_p$ where α_p are now i.i.d. standard Gaussian random variables.

Second, when there are periodic orbits of ψ in $\mathcal{M}$ of period less than the number of measurements M, it no longer holds true that $\operatorname{rank}(G_x - G_y) = M$ for x and/or y belonging to these periodic orbits. Since $r(G_x - G_y) \leq \operatorname{rank}(G_x - G_y)$, this means that we can no longer hope that the infimum soft-rank scales with M, thus rendering the stable embedding condition (e.g., (38)) mostly unattainable. Fortunately, we can circumvent this problem by separating out these periodic orbits from the analysis and providing conditions for them individually.

Third, we can remove the parasitic $\log(MP)$ in the infimum soft-rank condition (e.g., (38)) by using more sophisticated machinery. Recall that this term does not scale favorably when using monomials of degree up to $2M$ for the basis measurement functions. However, current analysis indicates that by removing this logarithm term, we pick up a quadratic term in the dimension of the attractor instead.

Fourth, we observe that this work is related the stable embedding of manifolds with block diagonal operators. To see this, write $F_\alpha(x) - F_\alpha(y)$ in the following way:

$$F_\alpha(x) - F_\alpha(y) = (G_x - G_y)\alpha = \begin{pmatrix} \alpha^T & & (0) \\ & \ddots & \\ (0) & & \alpha^T \end{pmatrix} (\bar{g}(x) - \bar{g}(y)) =: \Phi(\bar{g}(x) - \bar{g}(y)),$$

where $\Phi \in \mathbb{R}^{M \times MP}$ is a RBD matrix (see Chapter 3 and [107]). This correspondence to block diagonal matrices allows us to consider different measurement modalities, including multivariate time-series measurements and independent measurement functions at every time step T_s.

Fifth, while submanifolds in Euclidean spaces have many convenient analytic properties (e.g., the second fundamental form of the manifold which controls its curvature), many dissipative dynamical systems converge onto attractors that are not submanifolds.[21] In fact, stable embeddings provided by delay-coordinate maps can be extended to an arbitrary subset of Euclidean space that represents the attractor of a dissipative dynamical system. The easiest scenario arises when there is a *global enveloping manifold* $\widehat{\mathcal{M}}$ that subsumes the attractor $\mathcal{M}$. Roughly speaking, we say that $\widehat{\mathcal{M}}$ is a global enveloping manifold of an attractor $\mathcal{M}$ if $\mathcal{M} \subset \widehat{\mathcal{M}}$ and at every point $x \in \mathcal{M}$, $\mathcal{T}_x\mathcal{M} = \mathcal{T}_x\widehat{\mathcal{M}}$ (see [23, 105] for a more precise definition). Here, $\mathcal{T}_x\mathcal{M}$ is the *tangent cone* of $\mathcal{M}$ at x defined as the set of all secants converging to x (see [66] for more details). We call this a tangent cone instead of a tangent space since $\mathcal{M}$ is not a manifold. In such a scenario, Theorem 5.3.2 can be used to provide conditions for the stable embedding of the kernelized trajectories of the attractor via its global enveloping manifold. Alas, a counter-example in [23] shows that not all subsets of Euclidean space (and thus not all attractors of dynamical systems) can have a global enveloping submanifold. When the attractor does not have a global enveloping manifold, we endow the attractor $\mathcal{M}$ with a dimension, covering regularity, volume, and local smoothness properties that a submanifold intrinsically possesses. These are denoted by the box-counting dimension, box-counting regularity, and the tangent covering regularity. Recall that we have already defined the box-counting dimension of an attractor in Definition 5.1.1. Next, the box-counting covering regularity is a proxy for the volume of $\mathcal{M}$. Lastly, the tangent covering regularity has units of inverse length and can be thought of as a measure the curvature of $\mathcal{M}$ (and therefore related to the second fundamental form).

Finally, we showed in Section 5.3.3 a simulated dynamical system whose infimum

[21] Similarly, dynamical systems defined on manifolds may also converge onto a smaller, arbitrary subset of the manifold. These results apply to this case as well.

soft-rank scales favorably with M. Thus, the delay-coordinate map provides a stable embedding of this system's attractor for a reasonable number of delays M. For practical dynamical systems seen in the literature (e.g., the Lorenz system), their infimum soft-rank remains to be analyzed. This may entail converting the infimum soft-rank characterization of a dynamical system to more familiar characteristics like the Lyapunov exponents. Obtaining this conversion may allow us to differentiate systems that allow for stable embeddings by delay-coordinate maps for a reasonable number of delays M.

CHAPTER VI

MEMORY IN NEURAL NETWORKS

The ability of networked systems (including artificial or biological neuronal networks) to perform complex data processing tasks relies in part on their ability to encode signals from the recent past in the current network state. In this chapter, we use CS tools to study the ability of a particular network architecture (Echo State Networks) to stably store long input sequences. In particular, we show that such networks satisfy the RIP when the input sequences are compressible in certain bases and when the number of nodes scale linearly with the sparsity of the input sequence and logarithmically with its dimension.[1]

6.1 Background

When a networked system has recurrent (i.e., feedback) connectivity between the nodes, perturbations in the collective network states due to external inputs can persist long after the input has been removed. This persistence is a type of short term memory (STM) where the transient network state collectively retains information about past inputs. While there are many examples of networked systems with interesting memory characteristics (e.g., institutional memory in social networks), perhaps the most relevant example of this scenario is sequence memory in neural networks. Neural systems must hold sequences of stimulus information in STM for planning,

[1]This work was performed in collaboration with Adam S. Charles and Christopher J. Rozell. ASC and HLY contributed equally to this work. Specifically, ASC developed the initial problem formulation and ran substantial simulations, while HLY derived the appropriate RIP bounds. The full results of this work are in a submitted manuscript [36], while initial results were presented in [35, 142].

prediction and decision making. Researchers have postulated that the neural substrate for this type of STM is the current state of a richly connected recurrent neural network (sometimes called "reservoir computing" [79, 93]). In contrast, models of long-term memory likely rely on network connectivity changes (i.e., neural plasticity), and models of STM based on the notion of network attractors are primarily applicable to single patterns instead of input sequences.

Characterizing the fundamental limits of STM in networked systems is critical to understanding the computational abilities of these networks. For example, fundamental questions in this area include determining the effects on memory capacity of network size, connectivity patterns, and input statistics. Toward these questions, several researchers [62, 63, 78] have recently investigated network models of the form:

$$x[n] = f\left(Wx[n-1] + zs[n] + \widetilde{\epsilon}[n]\right), \tag{41}$$

where $x[n] \in \mathbb{R}^M$ are the network states at time n, W is the $(M \times M)$ recurrent (feedback) connectivity matrix, $s[n] \in \mathbb{R}$ is the input sequence at time n, z is the $(M \times 1)$ projection of the input into the network, $\widetilde{\epsilon}[n]$ is a potential network noise source, and $f : \mathbb{R}^M \to \mathbb{R}^M$ is a possible point-wise nonlinearity. The general idea is that if W is rich enough (often taken to be random connections), a single input will reverberate in the network, thereby creating a "memory" of the past input in the current network states. Figure 20 shows an illustration of how past inputs can drive the current network to different states, providing information to potentially recover the input history.

The STM capacity of the linear version of this network model (i.e., $f(x) = x$) has recently been extensively studied [62, 63, 78, 139], and this linear network will also be our focus in this chapter. While exact definitions of STM capacity vary, each approach in the literature attempts to quantify the amount of information in the current network state available for recovering a past input with some fidelity (e.g., the correlation between the input sequence and the recalled input estimate,

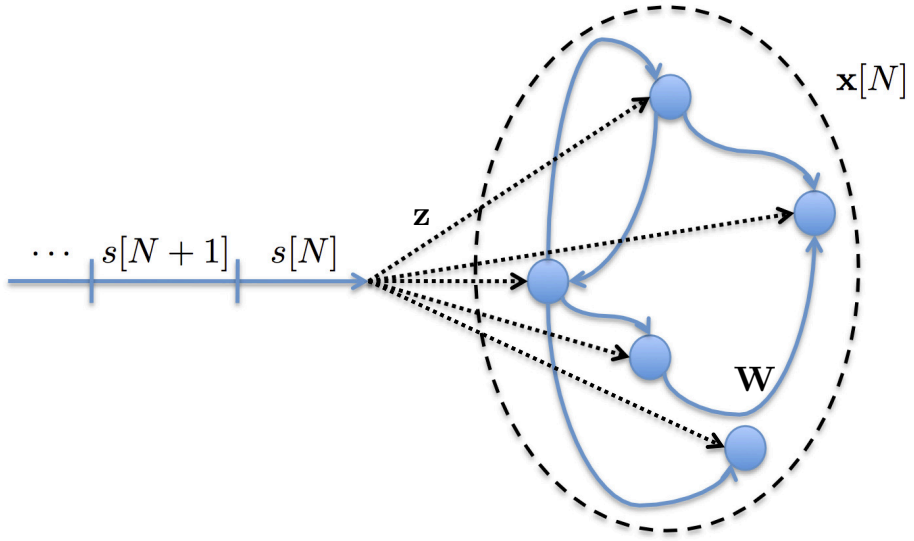


Figure 20: A pictorial description of the ESN showing the input at time N, $s[N]$, being fed by the feed-forward vector z into the reservoir of nodes $x[N]$ with connectivity pattern W.

Fisher information, etc.). These analyses rely on the stochastic nature of the input signal $s[n]$, with [78, 139] specifically assuming Gaussian statistics. These analyses derive STM capacity limits of $N \leq M$, meaning that the number of time samples significantly recoverable by the current network state is limited by the number of nodes in the network.

Recent work in computational neuroscience and signal processing has shown that many signals of interest have statistics that are strongly non-Gaussian, and that this additional structure can be highly leveraged for more efficient processing. In particular, sparsity-based signal models have recently been shown to be especially powerful. In the computational neuroscience literature, it has been demonstrated that sparsity models (coupled with the statistics of natural images) are sufficient to account for both the development of canonical response preferences [103] and the nonlinear response modulations [150] of neurons in the primary visual cortex. Furthermore, sparse signal encodings have been shown to increase the capacity of associative memory models [16].

In the signal processing literature, exploiting sparsity models for natural images underlies many state-of-the-art results in applications such as inverse problems (e.g., denoising, deblurring, superresolution, etc.) and computer vision [141]. Furthermore, the recent emergence of work in compressed sensing (CS) has shown that sparsity models can be used to make strong guarantees on signal recovery from highly undersampled measurement systems. In fact, using this sparsity model for the input statistics, Ganguli and Sompolinski [63] use an asymptotic statistical mechanics analysis on an approximation of the network dynamics in (41) to argue that orthogonal recurrent network structures can have STM capacities that exceed the number of network nodes.

In this work we prove strong guarantees on the STM capacity of the exact linear network dynamics given in (41) under various assumptions on the network structure. Our results leverage the established guarantees of the CS literature to characterize the network properties, as well as contributing to this literature by providing novel results on new measurement systems. The main contribution of our work is to provide rigorous, non-asymptotic recovery error bounds for sparse input sequences that show network STM capacities can be significantly higher than the number of the nodes in the network. Our analysis characterizes the impact on STM capacity of the input sparsity level and sparsity basis (e.g., sinusoids, wavelets, etc.), as well as the characteristics of the recurrent connectivity matrix. We provide both perfect recovery guarantees for finite inputs, as well as results on the recovery tradeoffs when the network has an infinitely long input sequence. The latter analysis highlights the fact that when the network has an infinitely long streaming input, the system has an optimal recovery length that balances errors due to omission and recall mistakes.

6.2 Network Dynamics as Compressed Sensing

In the network STM problem described above, we seek to understand the power of rich random connectivity structures for compressing a long input time series down to information contained in a small number of network nodes. These similarities to the basic CS problem lead us to seek analysis methods from the CS literature. To begin, we first assume that the input sequence $\{s[n]\}$ is sparse in the basis Ψ. Specifically, we assume that every length-N segment of the signal can be written using the basis Ψ with S non-zero coefficients.[2] Next, we write the network dynamics as a CS measurement operation. The recurrent dynamics of (41) (with $f(\cdot)$ being the identity) can be used to write the network state at time N in terms of the input signal and the iteratively applied connectivity matrix:

$$x[N] = \Phi s + \epsilon \tag{42}$$

where, Φ is an $M \times N$ matrix, the k^{th} column of Φ is $W^{k-1}z$, $s = [s[N], \ldots, s[1]]^T$, the initial state of the system is $x[0] = 0$, and ϵ is the node activity not accounted for by the input stimulus (e.g. the sum of network noise terms $\epsilon = \sum_{k=1}^{N} W^{N-k}\widetilde{\epsilon}[k]$).

Writing the network dynamics as a CS problem, we see that if Φ satisfies the RIP for the sparsity basis Ψ, the bounds given in Theorem 2.4.1 will establish strong guarantees on recovering s from the current network states $x[N]$ by solving the ℓ_1-optimization program (7). Note that this is a different approach from previous analyses on network STM using Gaussian statistics (as discussed in Section 6.1), where the recall process is a linear function of the network state. While the CS recovery process is nonlinear, there have been several recent proposals for network structures that efficiently solve the optimization (7) (e.g., [116]).

To leverage the CS recovery guarantees, we must establish the RIP conditions for Φ. We start from the definition of Φ in terms of the network connectivity matrix and

[2] We note that in some scenarios, it may make sense instead to define the sparsity to be a constant fraction η of the input length so that $S = \eta N$.

the feed-forward vector:

$$\Phi = \left[z \;\; | \;\; Wz \;\; | \;\; W^2 z \;\; | \;\; \cdots \;\; | \;\; W^{N-1} z \right].$$

Since we have a complete description of the connectivity of the network (and assuming it is full-rank with linearly independent eigenvectors), we can find the decomposition of the square matrix $W = UDU^{-1}$, where U is a matrix comprised of eigenvectors in each column and D is the diagonal matrix of eigenvalues. We can use this eigenvalue decomposition to rewrite the matrix Φ as

$$\Phi = U \left[\tilde{z} \;\; | \;\; D\tilde{z} \;\; | \;\; D^2 \tilde{z} \;\; | \;\; \cdots \;\; | \;\; D^{N-1} \tilde{z} \right],$$

where $\tilde{z} = U^{-1} z$. This setup can then be reorganized as

$$\Phi \;\; = \;\; U\widetilde{Z} \left[d^0 \;\; | \;\; d \;\; | \;\; d^2 \;\; | \;\; \cdots \;\; | \;\; d^{N-1} \right] = U\widetilde{Z}F$$

where $d = \text{diag}(D)$ is the column vector consisting of the eigenvalues of W, $\widetilde{Z} = \text{diag}(\tilde{z})$, the exponentiation of a vector is defined as the element-wise exponentiation, and F is a Vandermonde matrix with $F_{k,l} = d_k^l$ being the k-th eigenvalue of W raised to the l-th power.

Writing this decomposition for the 'measurement' process mapping the past input signal into the current network state highlights a few important issues. First, the matrix Φ is highly structured, with an exponentially changing dependency on signal components farther in the past. While the RIP conditioning of Φ depends on all of the matrices in the decomposition of (43), the conditioning of F is the most important and challenging component because it is compressive (i.e., not square). Second, we note the critical role that the eigenvalues of W play in the analysis. System stability requires that the eigenvalue magnitudes be less than or equal to 1, but smaller eigenvalues result in an exponential decay of a past input (i.e., the input will be "forgotten" by the network more quickly).

6.3 STM Capacity of Finite-length Inputs

We first consider the STM capacity of a network with finite-length inputs. While more restricted than the case of infinite-length inputs (which are considered in the next section), this scenario allows us to establish the most fundamental and general recovery guarantees. The basic setup is illustrated in Figure 21, where a linear encoding network (defined by W and z) is driven by an input sequence that is sparse in a wavelet basis, and recovery is performed by a nonlinear decoding network [116] that solves BPDN (7).

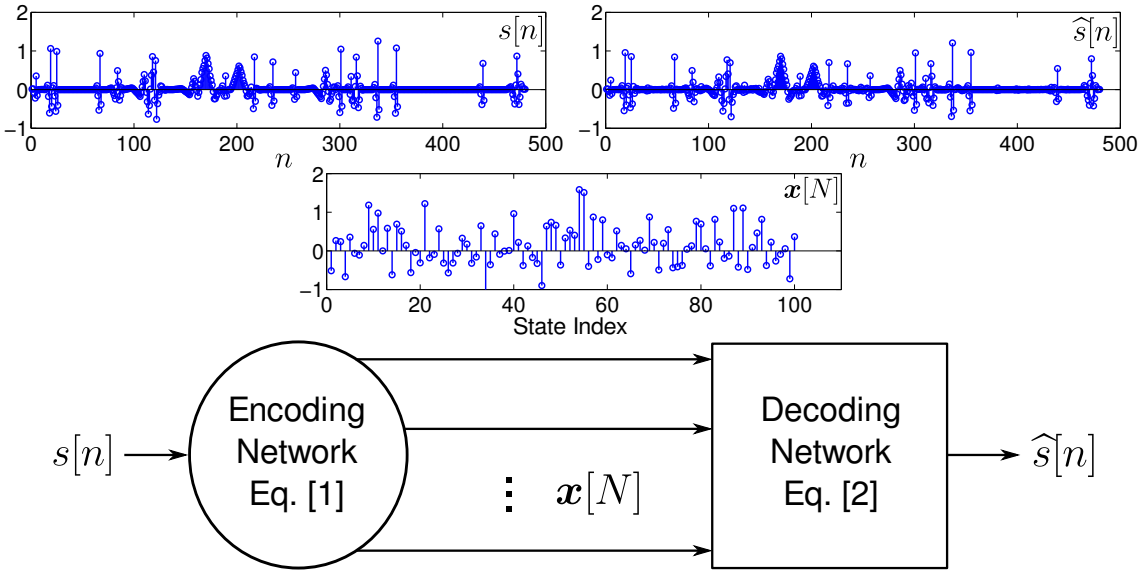


Figure 21: A length 480 stimulus pattern (left plot) that is sparse in a wavelet basis is fed into the encoding network defined by a random orthogonal matrix W and a feed-forward vector z. The 100 node values (center plot) are then used to recover the full stimulus pattern (right plot) using a decoding network which solves BPDN (7).

To leverage the CS recovery guarantees, we seek reasonable constructions of W that lead to matrix Φ satisfying the RIP. Because the structure of F is most limiting, we first assume structure for U and $\widetilde{Z}$ that preserves the conditioning properties of F in the definition of Φ. Specifically, as in [62, 63, 139] we choose W to be a random orthonormal matrix, assuring that the eigenvector matrix U has orthonormal

columns and therefore preserves the conditioning properties of F. Likewise, we choose the feed-forward vector z to be

$$z = \frac{1}{\sqrt{M}} U 1_M, \tag{43}$$

where 1_M is a vector of M ones. This choice assures that $\widetilde{Z}$ is the identity matrix scaled by $\sqrt{M}$.[3] Finally, we observe that the richest information preservation apparently arises for a real-valued W when its eigenvalues are complex, distinct in phase, have unit magnitude, and appear in complex conjugate pairs.

Under the construction above, the Vandermonde matrix F is a randomly subsampled Discrete Time Fourier Transform (DTFT). For this construction, the bounds from ℓ_1-recovery (see Theorem 2.4.1 in Section 2.4.1) hold as long as the number of nodes M satisfies the inequality

$$M \geq C \frac{S}{\delta^2} \mu^2 (\Psi) \log^4 (N) \log \left(\rho^{-1} \right), \tag{44}$$

where δ is the RIP conditioning of Φ, N is the length of the recovered input signal (i.e., the STM capacity where perfect recovery is possible in the noiseless case), S is the sparsity of the input sequence, C is a constant, $\mu(\Psi)$ quantifies the similarity between the sparsity basis Ψ and a randomly sampled DTFT, and ρ is a small predetermined probability of failure to satisfy RIP. The proof of this statement is given in Appendix F.1 and follows closely the approach in [111]. The quantity $\mu(\cdot)$ (known as the coherence) captures the largest inner product between the sparsity basis and the Fourier basis, and is calculated as:

$$\mu (\Psi) = \max_{n=1,\cdots,N} \sup_{t \in [0,2\pi]} \left| \sum_{m=0}^{N-1} \Psi_{m,n} e^{-jtm} \right|. \tag{45}$$

In the result above, the coherence is lower (therefore the STM capacity is higher) when the sparsity basis is more "different" from the Fourier basis.

[3] The factor of $\sqrt{M}$ is chosen to ensure that the columns of Φ have unit norm. This simplifies the recovery proofs, but has no bearing on the results.

The main observation of the result above is that for some values of S and $\mu(\Psi)$, it is possible to have STM capacities much greater than the size of the network (i.e., $N \gg M$). This is especially true when the input signal has a low sparsity S and a sparsity basis that has a low coherence with the Fourier basis. To illustrate that the analytic result above captures the STM gains and the dependence on the sparsity basis, Figure 22 shows the average recovery relative MSE (rMSE) in simulation from an input sequence of length N with M nodes. We use a plotting style similar to the phase transition diagrams of [55] where the average recovery rMSE is shown for each pair of variables under noise-free conditions.[4] The wedge between the dashed line ($M = N$) and the solid line (recovery error = 0.1%) in each plot shows the region where the STM capacity is significantly higher than M, and the input signal is essentially perfectly recovered.

6.4 STM Capacity of Infinite-length Inputs

After establishing the perfect recovery bounds for finite-length inputs in the previous section, we turn here to the more interesting case of a network that has received an input length beyond its STM capacity (or most generally, received infinite-length inputs). In contrast to the finite-length input case where favorable constructions for W used random unit-norm eigenvalues, such a construction here would be unstable as the network node values would grow without bound. In this case, we take W to have all eigenvalue magnitudes equal to $q < 1$ to ensure stability. The matrix constructions we consider in this section are otherwise identical to that described in the previous section.

In this scenario, the recurrent application of W in the system dynamics assures that each input perturbation will decay steadily until it has zero effect on the network state. While good for stability, this decay means that each input will slowly recede

[4]For equality-constrained recovery we use the l1 magic toolbox.

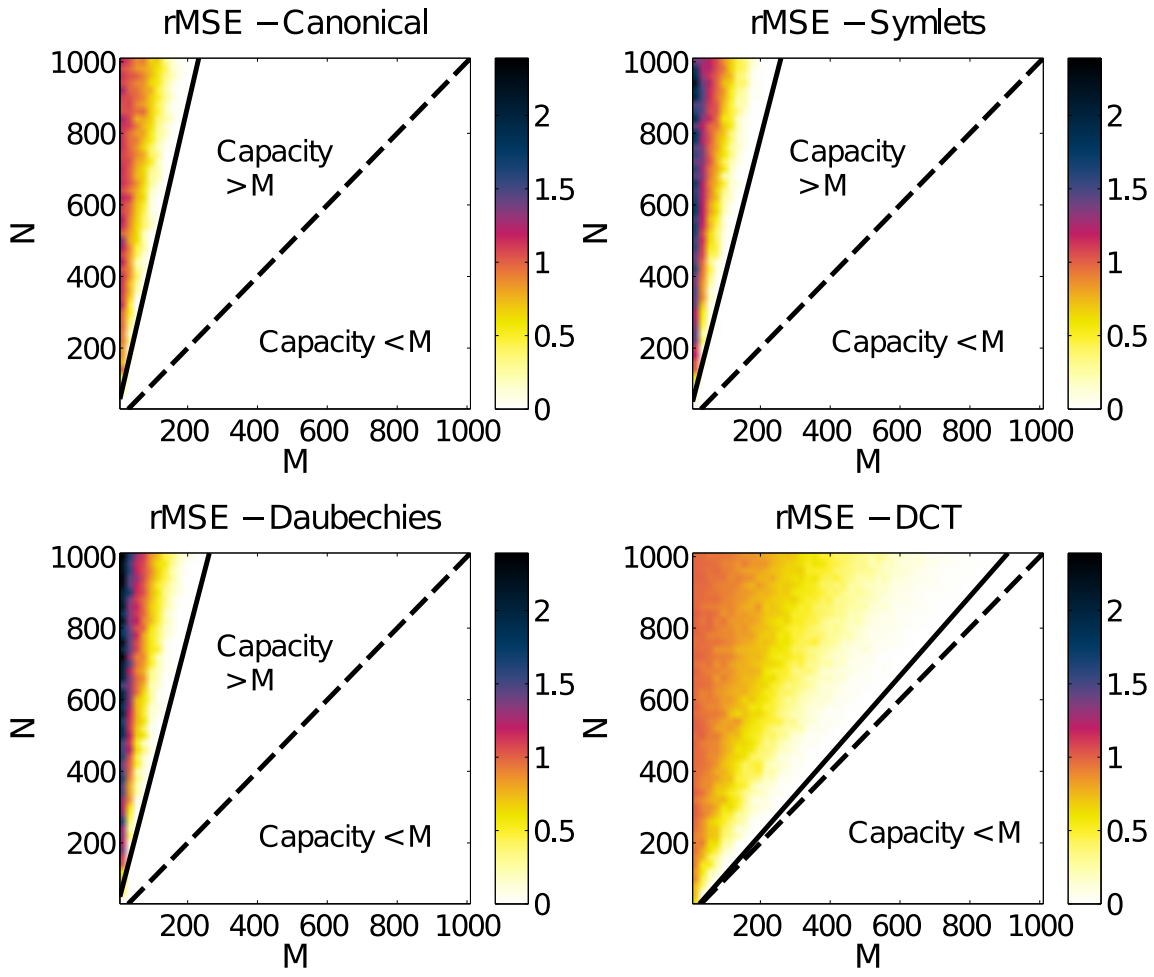


Figure 22: **Random orthogonal networks can have a STM capacity that exceeds the number of neurons. These plots depict the recovery relative mean square error (rMSE) for length-N input sequences from M network nodes where the input sequences are $S = \eta N$ sparse with $\eta = 0.05$. Recovery is near perfect (rMSE is machine precision) for large areas above the $N = M$ line for sequences sparse in the canonical basis as well as in various wavelet basis (shown here are 4 level decompositions in Symlet-3 wavelets and Daubechies-10 wavelets). For bases more coherent with the Fourier basis, recovery performance above $N = M$ suffers significantly, as shown in the phase plot using signals sparse in the discrete cosine transform (DCT).**

into the past. After some period of time, the network will no longer retain any useable memory of this event. In other words, *any* network can only hope to recover a proxy signal that accounts for the decay in the signal representation induced by the forgetting factor q. Specifically, we define this proxy signal to be Qs, where $Q = \text{diag}\left([1, q, q^2, \cdots]\right)$. Previous work [63, 78, 139] has characterized recoverability by using statistical arguments to quantify the correlation of the node values to each past input perturbation. In contrast, our approach is to provide a recovery bound for a network attempting to recover the N past samples of Qs. Note that in contrast to the previous section, the amount of time we attempt to recall, N, is now a parameter that can be varied.

Our technical approach to this problem comes from observing that inputs farther in the past than the attempted recovery length act as interference in the network when trying to recover more recent inputs. In other words, we can group older terms (i.e., from farther back than N time samples ago) with the noise term, resulting again in Φ being an M by N linear operator that can satisfy the RIP just as with finite-length inputs. In this case, after choosing the length of the memory to recover (N), the ℓ_1-recovery guarantees in Theorem 2.4.1 (see Section 2.4.1) hold when considering every input older than N contributing to the "noise" part of the bound.

Intuitively, we see that this approach implies the presence of an optimal value for the recovery length N. For example, choosing N too small means that there is useful signal information in the network that the system is not attempting to recover, and the system experiences errors by being too conservative and omitting potentially recoverable signal (i.e., an increase in the first term of (8) by counting too much signal as noise). On the other hand, choosing N too large means that the system is being too aggressive and is encountering recall errors by trying to recover inputs with little or no residual information remaining in the network activity (i.e., an increase in the second term of (8) from making the signal approximation worse by using the same

number of nodes for a longer signal length). Interestingly, this thought experiment indicates that if the goal is recovery of a short sequence, the optimal approach might be to recover a longer sequence (to account for more of the network activity and reduce recovery errors) and simply truncate the result.

Specifically, in the noiseless case where s is sparse in the canonical basis ($\mu(I) = 1$) with a maximum signal value $s_{\max}$, we can bound the first term of (8) using a geometric sum that depends on N, S and q. For a given scenario (i.e., a choice of q, S and the RIP conditioning of Φ), a network can support signal recovery up to a certain sparsity level S^*, given by:[5]

$$S^* = \frac{M\delta^2}{C\log^\gamma(N)}. \tag{46}$$

We can also bound the second term of (8) by the sum of the energy in the past N perturbations that are beyond this sparsity level S^*. Together these terms yield the bound on the recovery of the proxy signal:

$$\begin{aligned} \|Qs - Q\widehat{s}\|_2 \;\leq\; & \alpha s_{\max}\|U\|_2\left(\frac{q^N}{1-q}\right) \\ & + \frac{\beta s_{\max}}{\sqrt{\min\{S^*, S\}}}\left(\frac{q^{\min\{S^*,S\}} - q^S}{1-q}\right). \end{aligned} \tag{47}$$

The derivation of the above bound is detailed in Appendix F.2. The expression in (47), as shown in Figure 23, does have a minimum indicating that there is a value of N that achieves the best tradeoff between omission and recall errors. To test the validity of (47), we simulate recovery of different STM lengths and show the results in Figure 23. Note that the recovery MSE does match the qualitative behavior of the theoretical bound by reaching a minimum value, and that minimum is at $N > M$.

[5]While the proof techniques used here result in $\log^4$ dependency on N, the actual scaling law is conjectured to be $S\log(N)$ [118], and so here we leave the log factor γ intentionally ambiguous.

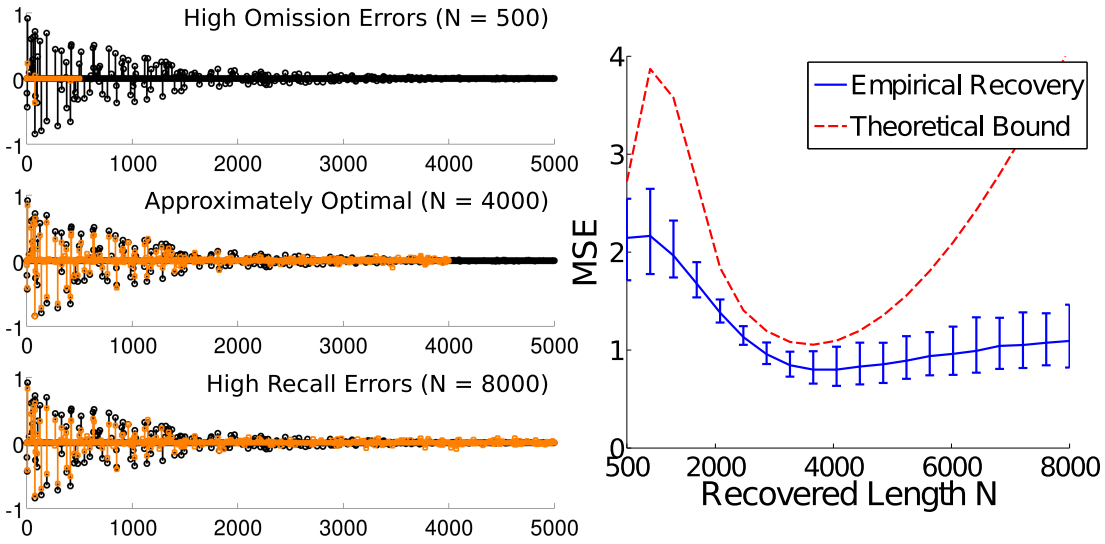


Figure 23: **The theoretical bounds for recovery of the past N perturbations to a network of size M has a minimum value at some optimal recovery length. This optimal value depends on the network size, the sparsity S, the decay rate q, and the RIP conditioning of Φ. Shown on the right is a simulation depicting the MSE for both the theoretical bound (red dashed line) and an empirical recovery for varying recovery lengths N. In this simulation $S = 400$, $q = 0.999$, $M = 500$. The error bars for the empirical curve show the maximum and minimum MSE. On the left we show recovery of a length 8000 signal when recovering the past 500 (top), 4000 (middle), and 8000 (bottom) most recent perturbations. As expected, at $N = 4000$ (approximately optimal) we recover the targeted values with the highest accuracy.**

6.5 Other Network Constructions

Our results in the previous sections focus on the case where W is orthogonal and z projects the signal evenly into all eigenvectors of W. When either W or z deviate from this structure the STM capacity of the network apparently decreases. In this section we revisit those specifications, considering alternate network structures allowed under these assumptions as well as the consequences of deviating from these assumptions in favor of other structural advantages for a system (e.g., wire length, etc.).

To begin, we consider the assumption of orthogonal network connectivity, where the eigenvalues have constant magnitude and the eigenvectors are orthonormal. Constructed in this way, U exactly preserves the conditioning of $\widetilde{Z}F$. While this construction may seem restrictive, orthogonal matrices are relatively simple to generate and encompass a number of distinct cases. For small networks, selecting the eigenvalues and choosing an orthonormal set of complex conjugate eigenvectors creates precisely these optimal properties. For larger matrices, the connectivity matrix can instead be constructed directly by choosing W at random and orthogonalizing the columns. Previous results on random matrices [51] guarantee that as the size of W increases, the eigenvalue probability density approaches the uniform distribution as desired.

While simple to generate in principle, the matrix constructions discussed above are all densely connected and may be impractical for many systems. However, many other special network topologies that may be more biophysically realistic (i.e., block diagonal connectivity matrices and small-world[6] networks [97]) can be constructed so that W still has orthonormal columns. For example, consider the case of a block diagonal connection matrix (illustrated in Figure 24), where many unconnected networks of at least two nodes each are driven by the same input stimulus and evolve

[6]Small-world structures are typically taken to be networks where small groups of neurons are densely connected amongst themselves, yet sparse connections to other groups reduces the maximum distance between any two nodes.

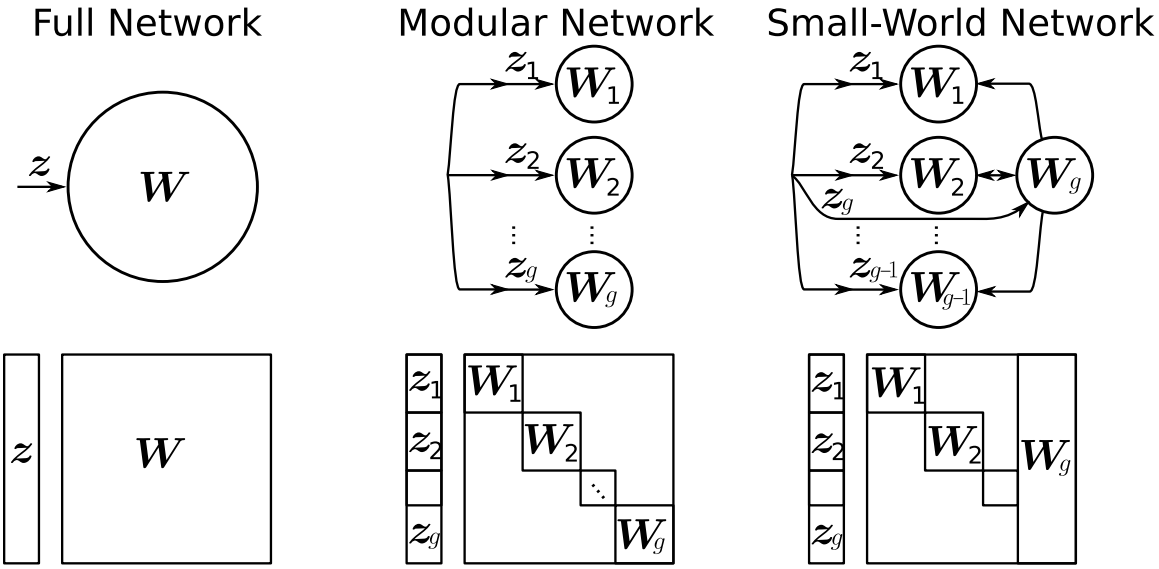


Figure 24: Possible network topologies which have orthogonal connectivity matrices. In the general case, all nodes are connected via non-symmetric connections. Modular topologies can still be orthogonal if each block is itself orthogonal. Small world topologies may also have orthogonal connectivity, especially when a few nodes are completely connected to a series of otherwise disjoint nodes.

separately. Such a structure lends itself to a modular framework, where more of these subnetworks can be recruited to recover input stimuli further in the past. Similarly, a small-world topology can be achieved by taking a few of the nodes in every group of the block diagonal case and allowing connections to all other neurons (either unidirectional or bidirectional connections). In these cases, the same eigenvalue distribution and eigenvector properties hold as the fully connected case, resulting in the same RIP guarantees (and therefore the same recovery guarantees) demonstrated earlier. We note here that one special case of orthogonal matrices not well-suited to the STM task are symmetric networks, where the strictly real-valued eigenvalues generates poor RIP conditioning for F.

There are a few variations in network construction that can also be considered. First, instead of the deterministic construction for z discussed in the earlier sections (analogous to [62] where z is optimized to maximize the SNR in the system), there

has also been interest in choosing z as i.i.d. random Gaussian values [62, 63]. In this case, it is also possible to show that Φ satisfies the RIP (with the same RIP conditioning δ as before) by paying an extra $\log(N)$ penalty in the number of measurements. The conditions for RIP to hold for this case is detailed in Appendix F.3. Second, instead of orthogonal connectivity matrices, there has also been interest in network constructions involving non-orthogonal connectivity matrices (perhaps for noise reduction purposes [62]). In this case, the eigenvector matrix U is no longer orthogonal and one can quantify how the conditioning of this matrix will affect the RIP conditioning of the system. A worst case analysis for the RIP conditioning of non-orthogonal matrices is provided in Appendix F.4. Note that in each of the two cases above, using these alternate constructions (which may be beneficial for other reasons) will have a detrimental effect on the RIP conditioning of the network.

6.6 Discussions

We have seen that the RIP property from the CS literature can provide a way to quantify the STM capacity in linear networks using rigorous, non-asymptotic bounds on the error of signal recovery. These results quantify the perfect recovery conditions for finite-length inputs while highlighting constructions of the network connectivity that achieve high STM capacities. This approach also allows us to bound the recovery error for infinite-length inputs, showing that there is an optimal recovery period that balances errors due to omission (i.e., not attempting to recall that portion of the signal) and recall mistakes. Of particular note is that this approach leverages the non-Gaussianity of the input statistics to show STM capacities that can be far above the number of nodes in the network. Specifically, the number of nodes needed in the network scales linearly with the sparsity level of the input and only logarithmically with the length of the signal being recovered.

We note that this approach is fundamentally different from past work that assumes

the input recovery process is a linear function of the current network state [78, 139]. Such an approach can only be optimal for Gaussian input statistics, and the fundamental results in the CS literature demonstrate clearly that the sparse signal models are only powerful when combined with a nonlinear inference process for signal recovery. This encoding and decoding process is asymmetric, making the storage of memories computationally efficient and robust by spreading the information across many nodes using linear computations. The nonlinear computation necessary for minimizing the ℓ_1-program only needs to be performed during recall, and this optimization can be performed in other (nonlinear) network architectures that are both efficient and biologically plausible [116]. We note also that despite the nonlinearity of the recovery process, the fundamental results of the CS literature also guarantee that the recovery process is stable and robust. With access to only a subset of nodes (due to failures or communication constraints), recovery generally degrades gracefully by still achieving the best possible approximation of the signal using fewer coefficients.

CHAPTER VII

CONCLUSION AND FUTURE DIRECTIONS

In this thesis, we analyzed conditions under which measurements taken of low-dimensional signals can be a stable embedding of the corresponding signal family. Our contributions include:

1. deriving the concentration of measure inequalities of random block diagonal matrices and showing that these matrices can satisfy the RIP with measurement rates that scale with a coherence property of the sparsity basis (and naturally with the sparsity and logarithmically with the ambient dimension);

2. showing that all operators satisfying the RIP can be used for stable manifold embeddings, and giving several examples of such operators that provide fast dimensionality reduction of manifold-modeled signals;

3. giving conditions under which delay-coordinate maps can be stable embeddings of the dynamic signals they are observing;

4. and showing that a recurrent network of neurons can stably hold memory of past (sparse) inputs, and this memory capacity can greatly exceed the number of nodes in the network.

As alluded to in the introduction, the measurement systems studied in this thesis are systems having various constraints. These constraints are either architectural (e.g., block diagonal, delay coordinate, or recurrent) or for efficiency (e.g., efficient transforms for stable manifold embeddings). Analyzing the resulting structured matrices that represent these systems requires more sophistical probabilistic and geometric machinery than their unstructured counterparts. However, this machinery could

be the reason why there is an increase in the measurement rates required for these systems to provide a stable embedding (e.g., requiring $S \log^{\alpha} N$ for $\alpha > 1$ instead of $S \log N$ measurements to satisfy RIP-(S, δ)). Nontheless, progress in the probability theory in Banach spaces is rapidly improving the tools necessary to improve these measurement rates. To give an example, measurement rates for a partial circulant matrix to satisfy the RIP-(S, δ) have reduced from the order of $(S \log N)^{3/2}$ [112] to $S \log^2 S \log^2 N$ [83]. The same methods used to affect this improvement have also been used to improve the measurement rates of RBD matrices by the same order.

A second drawback of constrained measurement systems is that their measurement rates are typically not universal. For example, when studying the memory capacity of recurrent neural networks, we have shown that such networks satisfy the RIP-(S, δ) with a measurement rate that depends on the coherence of the sparsity basis. Therefore, recurrent networks have good memory capacity for inputs sparse in certain bases but not others. We have also seen this non-universality in the measurement rates of delay-coordinate maps and block diagonal matrices. Nonetheless, we have seen that all non-universal RIP matrices can be made universal for any signal family just by randomizing their column signs. However, architectural constraints on various measurement processes may prevent the widespread use of this result. For example, it is not practical to modulate every (finite length) input sequence to a recurrent network with the same Rademacher sequence due to the streaming nature of the inputs.

Stable embedding guarantees by constrainted measurement systems are made possible by low-dimensional geometric models imposed on the signals they are designed to measure. Research in the signal processing community continues to uncover low-dimensional geometric signal models in an increasing number of applications (e.g., [30, 32, 128]). For example, the problem of restoring the phase of a signal where we have only captured the magnitude of its (random) measurements can be

posed as a low-rank matrix recovery problem [30] (in fact, the signal is modeled as a rank-one matrix). Also, the work by Chandrasekaran et al. [32] has shown that for general geometric signal sets (instead of sets based on sparsity), an atomic norm minimization (instead of BPDN in CS) may yield the correct solution of a underdetermined inverse problem as long as a specific nullspace property expressed with this norm is fulfilled. As initial theoretical analyses with these low-dimensional geometric models typically uses unstructured Gaussian measurement operators, the need to study constrained measurement systems of such signals grows in lockstep.

To be fair, stable embeddings are not the only criteria that researchers have used to establish recoverability guarantees of low-dimensional signals. First, stable embeddings based on other norms can be used as a sufficient condition for recovery using BPDN. For example, the ℓ_1-norm has been used as a distance metric in the RIP statement (6) instead of the typical ℓ_2-norm [67]. Second, some recovery results in the literature are non-uniform in nature [10, 27, 110, 111]. Non-uniformity means that for every vector x being measured, its recoverability is only guaranteed for a *separate* draw of a random measurement matrix Φ. Contrast this with the uniform RIP-based recovery result in Theorem 2.4.1, where for a *single* draw of the matrix Φ, we can guarantee the recoverability of all sparse vectors x. While non-uniform recovery results usually allow for better measurement rates M (e.g., for circulant matrices, compare [110] and [112]), stable embeddings remain relevant for a few reasons. First, in many applications, we would require a measurement system to work for *all* signals in a family once it is built. Second, the RIP gives a clean geometric interpretation (i.e., distance preservation) that will be useful for many signal processing/data inference algorithms other than signal recovery.

To conclude this thesis, we will discuss the remaining open problems for each of our contributions above, make links to other research areas, and describe possible future directions for research.

7.1 Random Block Diagonal Matrices

As we have discussed, block diagonal matrices represent measurement systems that are constrained to make only local observations (in time or in space). Indeed, many real-world systems have this locality constraint, and they deserve a more in-depth study with the tools we have developed for block diagonal matrices. These systems include distributed sensor networks (in particular MIMO radars [70, 89]), streaming video [21, 22], dynamic CS (more on dynamic CS will be discussed later), and the human visual system [77].

In our study, the non-zero blocks in the block diagonal matrices are *unstructured*, i.e., they are filled with i.i.d. subgaussian random variables. In certain situations, it might be necessary to consider *structured* random blocks. For example in a distributed radar system, (random) radar pulses that are sent out by different transmitters are convolved with the sparse target scene before being captured by a distributed array of receiver. Such a measurement process is best represented by a block diagonal matrix with random Toeplitz matrices on its block.

In fact, the locality constraint of block diagonal matrices is also linked to work on dynamic CS. In dynamic CS, the signal of interest $x_n \in \mathbb{R}^N$ are the states of a (linear) dynamical system:

$$x_n = A(x_{n-1}) + v_n,$$

where $A \in \mathbb{R}^{N \times N}$ and v_n is a noise process. At each time instance n, we only get to make compressive measurements of the state x_n, i.e., we measure

$$y_n = C_n x_n + w_n,$$

where $C_n \in \mathbb{R}^{M \times N}$ is a measurement matrix (which may be the same for all n) and w_n is additive measurement noise. To determine if we can recover the states x_n, an

often-used metric is the rank of the observability matrix O_J, which can be written as

$$\bar{y}_J = \begin{pmatrix} y_1 \\ \vdots \\ y_J \end{pmatrix} = \begin{pmatrix} C_1 A^0 \\ \vdots \\ C_J A^{J-1} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_J \end{pmatrix} = O_J \bar{x}_J.$$

Following the main themes of this thesis, the knowledge of just the rank of O_J is insufficient under noisy conditions. The study of O_J beyond just its rank has been studied in [138] where the authors expanded the observability matrix into a block diagonal matrix multiplied by a tall matrix of the powers of A:

$$O_J = \begin{pmatrix} C_1 & & \\ & \ddots & \\ & & C_J \end{pmatrix} \begin{pmatrix} A^0 \\ \vdots \\ A^{J-1} \end{pmatrix}.$$

The authors then did an empirical study of the recoverability of the states of the system using our concentration of measure analysis. It remains to be seen if our RIP analysis can help provide theoretical guarantees on recovery. The main difficulty will be to express

$$\begin{pmatrix} A^0 \\ \vdots \\ A^{J-1} \end{pmatrix} \bar{x}_J$$

as $U\bar{\alpha}$, where U is a sparsity basis of sufficiently small coherence and $\bar{\alpha}$ are coefficients of sufficiently low sparsity. This observability study can also be a pre-requisite for Kalman-filter like algorithms that try to recover the system states from compressive measurements, usually by exploiting the sparsity structure of the signals [8, 33, 34]. Finally, block diagonal matrices have also come up in the study of Multiple Measurement Vectors (MMV) [151].

7.2 Stable Manifold and Stable Takens' Embedding

Delay-coordinate maps can provide stable embeddings of the dynamic signals they are measuring, but the number of measurements may exceed the ambient dimensions of the signal. A common approach for (further) dimensionality reduction in delay-coordinate maps is through filtering the time-series output of the dynamical system. Existing theoretical guarantees for the resulting *filtered delay-coordinate maps* provide just an embedding of the attractor manifold in the *filtered reconstruction space* (just as how Takens' embedding provide just an embedding of the attractor manifold in the reconstruction space). As usual, without some form of stability, an embedding can be vulnerable to noise. By using our result on stable manifold embeddings, we can show that many efficient filtering operations can be devised such that the filtered attractor manifold has almost the same geometry as the attractor manifold in the reconstruction space. Figure 25 shows a block diagram of this filtering process, where we make explicit the pre-processing step of randomizing the time series by a Rademacher sequence.

On the other hand, we have only quantified the robustness of a manifold representation with stable embeddings, which means that the geometry of the ambient manifold is preserved in the measurement/reconstruction space. In some situations, it may be helpful to change the geometry of the manifold to suit other processing algorithms downstream. For example, it is heuristically argued that delay-coordinate maps that maximally unfold the attractor manifold of the input dynamical system are beneficial for downstream time-series prediction algorithms [31, 61]. This unfolding typically is done by a judicious choice of sampling time T_s and/or number of measurements M [61, 81, 95, 135]. For more generic types of dimensionality reduction operators (other than delay embeddings), there has also been work on unfolding of the input manifold under the rubric of manifold learning [115, 129], and more recently, diffusion maps [40, 41]. It remains an open question if our notions of stable

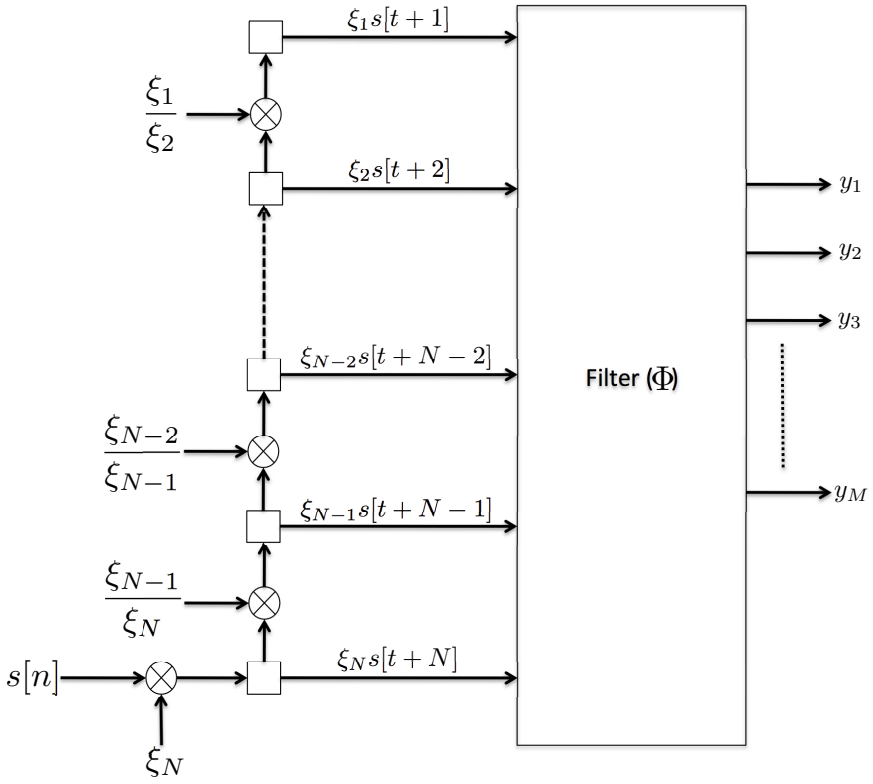


Figure 25: Block diagram showing a stable filtering process that preserves the geometry of the attractor manifold in the reconstruction space.

embedding and the corresponding tools to show such embeddings can be extended to include these manifold unfolding notions that are useful in these other context.

7.3 Network Memory

We showed that a recurrent network of neurons can store memory of a long input sequence in its transient states. But, modeling the time-series input as compressible in a certain basis does not take into account the dynamical correlations between the time-series values. What if the time-series input is as a one-dimensional function of a (possibly high-dimensional) dynamical system whose states have converged onto a low-dimensional attractor? Our results on stable Takens' embedding (or Takens' original result) state that when the time-series values are stacked up into a vector of

sufficient length, the resulting delay-coordinate map can be a one-to-one embedding or stable embedding of the system attractor. This means that the input sequence into an ESN can be a representation of the system attractor, and if it is further shown that the ESN can retain memory of its input sequence, then the transient network state of the ESN can also be a representation of the system attractor. It thus remains to show that the ESN can indeed retain memory of such dynamically modeled input sequences (which is not evident from our previous work which only deals with compressible inputs).

The results in Chapter 6 have also made several simplifying assumptions about the network connectivity and network topology. First, for analytical simplicity, we have assumed a linear recurrent network, meaning that the activation function from one time step to another is the identity. Nonlinearities in the activation function may be useful as a means of projecting the input signals into a (higher-dimensional) *kernel* space, as is evidenced by our work on stable Takens' embedding with nonlinear observation functions. This kernel space embedding may be useful for downstream algorithms that we wish to perform on the network states [90, 109]. Second, we have dealt mainly with an orthogonal connectivity matrix, and we gave brief, worst-case-scenario treatment to the situation where the connectivity matrix has singular values that are not all ones. However, non-orthogonal matrices can be useful for noise-reduction purposes [62], and non-orthogonal matrices can arise from typical (and empirically desirable) ESN constructions which have sparsely and randomly connected nodes [90].

ESNs [90] can perform complex data processing tasks while necessitating only relatively simple learning rules. To train an ESN to perform a specific computing task, the ESN is typically presented with a batch of training input sequences $s^k_{\text{train}} = \left[s^k[1], \cdots, s^k[N]\right]^T \in \mathbb{R}^N$ for $k = 1, \cdots, K_{\text{train}}$ and an output function $W_{\text{out}} : \mathbb{R}^M \to \mathbb{R}$ is then learned on the resulting network state $x^k[N] \in \mathbb{R}^M$ so that the outputs $y^k =$

$W_{\text{out}}(x^k[N])$ match their corresponding desired outputs $y^k_{\text{train}} \in \mathbb{R}$. This learning is typically done with machine learning tools such as SVM regression [123] or a simple linear regression while trying to minimize some error criterion between the output y^k and y^k_{train}, e.g., the ℓ_2-error $\left\|y^k - y^k_{\text{train}}\right\|_2$.

The success of reservoir computing, as the process discussed above is called, is thus dependent on two separable tasks. First, through the choice of the connectivity matrix W and the feedforward vector z, the ESN has be to designed such that its nodes can be a "good" representation of the input sequences feeding the network. The second task is the machine-learning task of learning an output function W_{out} so that the output value y^k is as close as possible (depending on the minimizing function chosen) to the expected output y^k_{train}. Invariably, these two tasks are intertwined such that the quality of the representation afforded by the first task will affect the success of learning the output function W_{out} in the second task. A robust representation of the input sequences in the network nodes will result in good learning capabilities for the output function W_{out}.

In this thesis, we have expressed the robustness of a representation as the near-isometry (stable embedding) of the input signal family in the representation space, and the deviation from isometry is usually measured in the ℓ_2-norm. Work in the CS community has also shown that considering other norms (or even mixed norms) may yield more favorable results in specific scenarios. Two examples are given earlier in this thesis. First in [67], recovery and stability guarantees are expressed using the ℓ_1-norm rather than with the ℓ_2-norm. Second, atomic norm minimization has been shown to yield correct solutions to underdetermined inverse problems of general geometric signal sets [32]. This atomic norm generalization may mean that near-isometry with different norms may be more appropriate with signal family geometries that are not based on sparsity. More generally, different notions of stability besides near-isometries may be helpful for different output function learning algorithms that

will be used downstream. For example, an existing notion that is currently used in the reservoir computing community is called the *kernel quality* [92]. Suppose $x^1[N], \cdots, x^K[N]$ are the node representation of K distinct input sequences $s^1, \cdots, s^K$ with $s^k = \left[s^k[1], \cdots, s^k[N]\right]$. The kernel quality is then the rank of the $M \times K$ matrix X whose columns comprise the K node representations $x^k[N]$. If the rank is K ($K \leq M$), then the reservoir is said to have a *linear separation property*, meaning that all the K signals are distinguishable. However, the linear separation property is not robust under noise (just like an embedding is not robust under noise). Thus, an improved measure to the kernel quality could instead be the singular value spread of the matrix X. First, if the singular values are all non-zero, then the reservoir has a linear separation property. On top of that, if the spread of singular values is small, then the node representations are maximally spread out in $\mathbb{R}^M$, i.e., they are of the same length and are orthogonal to one another. This implies that small perturbation of the node values by noise will have minimal impact on the separability of these representations. While the singular value spread of X can be calculated for a given (finite) set of input sequences and an ESN, it remains to be shown if we can provide a theoretical analysis to what properties of any ESN can result in a small spread of the singular values of X by drawing from recent tools in probability theory that has been used successfully in CS. In addition, it is also not clear how to extend this measure to typical cases where the signal family has infinite cardinality.

APPENDIX A

TOOLS OF THE TRADE

Before we begin our proofs, let us consolidate several lemmas that will be used throughout the appendices.

A.1 Probabilistic Tools

A random variable Z is *symmetric* if Z and $-Z$ has the same distribution. The symmetrization technique, shown in the following lemma, converts any centered random variable (or a sequence of centered random variables) into a symmetric random variable, and is a useful first step in many proofs.

Lemma A.1.1 (Symmetrization)**.** *Suppose that for $m \in [M]$, Z_m is sequence of independent Banach space vectors equipped with a norm $\|\cdot\|$. Then for $1 \leq p < \infty$,*

$$\left(\mathbb{E}\left\{\left\|\sum_{m=1}^{M} Z_m - \mathbb{E}\left\{Z_m\right\}\right\|^p\right\}\right)^{1/p} \leq 2\left(\mathbb{E}\left\{\left\|\sum_{m=1}^{M} \xi_m Z_m\right\|^p\right\}\right)^{1/p},$$

where $\{\xi_m\}$ is a Rademacher sequence independent of $\{Z_m\}$.

For most of the proofs below, we are interested in characterizing the tail bound of a random variable. The following lemma says that by knowing the moments of a random variable, we can easily get to its tail bound:

Lemma A.1.2 (Moments to Tail Bounds, adapted from Proposition 6.5 of [111])**.** *Suppose Z is a random variable satisfying*

$$\left(\mathbb{E}\left\{|Z|^p\right\}\right)^{1/p} \leq \alpha\beta^{1/p}p^{1/\gamma},$$

for all $p \in [p_0,\ p_1]$, and for constants $\alpha, \beta, \gamma, p_0, p_1$. Then, for all $u \in [p_0^{1/\gamma},\ p_1^{1/\gamma}]$,

$$\mathbb{P}\big\{|Z| \geq e^{1/\gamma}\alpha u\big\} \leq \beta e^{-u^\gamma/\gamma}.$$

In fact, there's an equivalence relation between the moments and the tail bounds of a random variable via an integral.

Lemma A.1.3 (Moments to Tail Bounds Equality)**.** *[111] For any random variable Z and any $p > 0$,*

$$\mathbb{E}\left\{|Z|^p\right\} \;=\; \int_0^\infty \mathbb{P}\{|Z| > t\}\, t^{p-1} dt.$$

The moments of a random variable is also a metric on the space of random variables, therefore, we have the triangle inequality of moments norm

$$\left(\mathbb{E}\left\{|X+Y|^p\right\}\right)^{1/p} \leq \left(\mathbb{E}\left\{|X|^p\right\}\right)^{1/p} + \left(\mathbb{E}\left\{|Y|^p\right\}\right)^{1/p},$$

for any $p > 0$ and X, Y are random variables.

When dealing with a function of a random variable, Jensen's inequality is an indispensable tool.

Lemma A.1.4 (Jensen's inequality)**.** *Let Z be a random variable and f a convex function, then $f(\mathbb{E}\{Z\}) \leq \mathbb{E}\{f(Z)\}$. If f is a concave function, then $\mathbb{E}\{f(Z)\} \leq f(\mathbb{E}\{Z\})$.*

Most of the proofs below deal with subgaussian random variables, and usually we are interested in a sum of squares of these subgaussian random variables. Bernstein's inequality shown in the lemma below gives a tight bound on the tail bound of this sum.

Lemma A.1.5. *[136] Let $Z_1, \cdots, Z_L$ be independent subgaussian random variables with subgaussian norms $\|Z_i\|_{\psi_2}$ for all $i = 1, \cdots, L$. Let $T = \max_i \|Z_i\|_{\psi_2}^2$. Then for every $t \geq 0$ and every $a = (a_1, \cdots, a_L) \in \mathbb{R}^L$, we have*

$$\mathbb{P}\left\{\left|\sum_{i=1}^{L} a_i \left(Z_i^2 - \mathbb{E}\left\{Z_i^2\right\}\right)\right| \geq t\right\} \leq 2\exp\left(-C_1 \min\left\{\frac{t^2}{16T^2\|a\|_2^2}, \frac{t}{4T\|a\|_\infty}\right\}\right),$$

where $C_1 > 0$ is an absolute constant.

Finally, the following lemma gives us estimates for a sum of random matrices:

Lemma A.1.6 (Weak Khintchine's Inequality)**.** *Let $\{A_l\}$ for $l \in [L]$ be a sequence of matrices of the same dimension and rank of at most $J \gtrsim 1$. Let ξ_l for $l \in [L]$ be a Rademacher sequence. Then, we have*

$$\mathbb{E}\left\{\left\|\sum_{l\in[L]}\xi_l A_l\right\|_2\right\} \lesssim \sqrt{\log J}\left(\sum_{l\in[L]}\|A_l\|_2^2\right)^{1/2}.$$

Proof. For $1 \leq p \leq \infty$, the Schatten norm of order p of a matrix A is denoted by $\|A\|_{S_p}$ and is defined as

$$\|A\|_{S_p} := \|\sigma(A)\|_p,$$

where $\sigma(A)$ is the vector formed by the singular values of A. Observe that $\|A\|_{S_\infty} = \|A\|_2$ and $\|A\|_{S_2} = \|A\|_F$. Schatten norms possess the following useful property that mirrors Euclidean norms:

$$\operatorname{rank}(A)^{\frac{1}{p}-\frac{1}{q}}\|A\|_{S_q} \leq \|A\|_{S_p} \leq \|A\|_{S_q}, \tag{48}$$

where $1 \leq q \leq p$.

In the literature, we have the following bounds for the Schatten norm of order p of a sum of random matrices that holds for every $2 \leq p < \infty$ [111, Theorem 6.14]:

$$\mathbb{E}_p\left\{\left\|\sum_{l\in[L]}\xi_l A_l\right\|_{S_p}\right\} \lesssim \sqrt{p}\max\left\{\left\|\left(\sum_{l\in[L]}A_l A_l^*\right)^{1/2}\right\|_{S_p}, \left\|\left(\sum_{l\in[L]}A_l^* A_l\right)^{1/2}\right\|_{S_p}\right\}.$$

However, Lemma A.1.6 is about the Schatten norm of order $p = \infty$, and therefore the inequality above does not apply to our problem. As such, we need a more detailed argument here, and this is achieved by following the footsteps of [136]. From (48), with $p = \infty$ and $q = \log J$, we have

$$e^{-1}\|A\|_{S_{\log J}} \leq \|A\|_2 \leq \|A\|_{S_{\log J}}.$$

This equivalence, in combination with the fact that moment norms are increasing in p, allows us to write

$$\begin{aligned}
\mathbb{E}\left\{\left\|\sum_{l\in[L]}\xi_l A_l\right\|_2\right\} &\leq \mathbb{E}\left\{\left\|\sum_{l\in[L]}\xi_l A_l\right\|_{S_{\log J}}\right\} \\
&\leq \mathbb{E}_{\log J}\left\{\left\|\sum_{l\in[L]}\xi_l A_l\right\|_{S_{\log J}}\right\} \\
&\lesssim \sqrt{\log J}\max\left\{\left\|\left(\sum_{l\in[L]}A_l A_l^*\right)^{1/2}\right\|_{S_{\log J}}, \left\|\left(\sum_{l\in[L]}A_l^* A_l\right)^{1/2}\right\|_{S_{\log J}}\right\} \\
&\leq e\sqrt{\log J}\max\left\{\left\|\left(\sum_{l\in[L]}A_l A_l^*\right)^{1/2}\right\|_2, \left\|\left(\sum_{l\in[L]}A_l^* A_l\right)^{1/2}\right\|_2\right\} \\
&= e\sqrt{\log J}\max\left\{\left\|\sum_{l\in[L]}A_l A_l^*\right\|_2^{1/2}, \left\|\sum_{l\in[L]}A_l^* A_l\right\|_2^{1/2}\right\} \\
&\leq e\sqrt{\log J}\max\left\{\left(\sum_{l\in[L]}\|A_l\|_2^2\right)^{1/2}, \left(\sum_{l\in[L]}\|A_l\|_2^2\right)^{1/2}\right\} \\
&= e\sqrt{\log J}\left(\sum_{l\in[L]}\|A_l\|_2^2\right)^{1/2},
\end{aligned}$$

as claimed. We assumed above that $J \geq e$ to produce the first line and the second to the last line above uses the triangle inequality and that $\|AA^*\|_2 = \|A\|_2^2$ for any matrix A. We remark that had we stopped at the fifth line above, we would have ended with the stronger original non-commutative Khintchine inequality for the spectral norm. However, the weaker and more amenable bound given in the last line suffices for our purposes, and this completes the proof of Lemma A.1.6. □

APPENDIX B

MANIFOLD PROPERTIES PROOFS

In this appendix, we prove the two lemmas regarding the consequences of the uniform boundedness of the second fundamental form (Lemma 2.5.1) and the geodesic regularity (Lemma 2.5.2).

B.1 Proof of Lemma 2.5.1

The proof of Lemma 2.5.1 follows closely Section 6 of [102]. Before we begin the proof, we note that we can quantify the rate of change of a parallel transport along a unit speed geodesic curve with the uniform boundedness of the second fundamental form. To define what we mean by a parallel transport, first let $\gamma(t)$ denote a unit speed geodesic curve joining p and q in $\mathcal{M}$ such that $\gamma(0) = p$, $\gamma(\mu) = q$, and $\|\gamma'(t)\|_2 = 1$ for all $t \in [0,\ \mu]$ with $\mu := d_{\mathcal{M}}(p, q)$ and $\gamma'(t) \in \mathcal{T}_{\gamma(t)}\mathcal{M}$, and let $v_0 \in \mathcal{T}_{\gamma(0)}\mathcal{M}$. We say that $v(t)$ is the *parallel transport* of v_0 along the unit speed geodesic curve γ if $v(0) = v_0$, $\|v(t)\|_2 = \|v_0\|_2$, $v(t) \in \mathcal{T}_{\gamma(t)}\mathcal{M}$, and $\langle v(t), \gamma'(t)\rangle = \langle v_0, \gamma'(0)\rangle$ (thus the name "parallel"). The following lemma says that if the submanifold $\mathcal{M}$ has bounded second fundamental form, then the length of $v'(t)$ is bounded.

Lemma B.1.1. *Let p, q be two distinct points on a submanifold $\mathcal{M} \subset \mathbb{R}^N$ whose second fundamental form is uniformly bounded by $\frac{1}{\tau}$. Denote by $\gamma(t)$ the unit speed parametrization of the geodesic path joining p and q. Suppose we have a vector $v_0 \in \mathcal{T}_{\gamma(0)}\mathcal{M}$ and let $v(t)$ be the parallel transport of v_0 along the curve γ. Then, $\|v'(t)\|_2 \leq \frac{\|v_0\|_2}{\tau}$ for all $t \in [0,\ \mu]$.*

Proof. From our discussion on connections on manifolds representing differentiation along a curve in Section 2.5.1, we have $v' := \nabla_{\gamma'} v$. It can be shown that for a parallel

transport, $(\nabla_{\gamma'} v)^T = 0$, i.e., the projection of the derivative of v onto the tangent space of the manifold is zero [52]. Thus, $v' = (\nabla_{\gamma'} v)^\perp = B(\gamma', v)$. Since the second fundamental form of $\mathcal{M}$ is uniformly bounded, we have that for all $p \in \mathcal{M}$ and for all $\eta \in (\mathcal{T}_p\mathcal{M})^\perp$, $\|S_\eta\|_2 \leq \frac{1}{\tau}$. Therefore, for $\eta := \frac{v'(t)}{\|v'(t)\|_2}$,

$$\|v'(t)\|_2 = \langle v'(t), \eta \rangle = \langle B(\gamma', v)_{\gamma(t)}, \eta \rangle = \langle \gamma'(t), S_\eta v(t) \rangle \leq \|\gamma'(t)\|_2 \|S_\eta\|_2 \|v(t)\|_2,$$

where substituting $\|\gamma'(t)\|_2 = 1$, $\|S_\eta\|_2 \leq \frac{1}{\tau}$, and $\|v(t)\|_2 = \|v_0\|_2$ concludes the proof. □

With this lemma, we can prove Lemma 2.5.1 concerning properties of manifolds induced by the boundedness of the second fundamental form.

Proof of Lemma 2.5.1. The first part of the lemma is a simple consequence of Lemma B.1.1. Let $\gamma(t)$ be the unit speed parameterization of the geodesic curve between p and q as before. We want to show that $\|\gamma''(t)\|_2 \leq \frac{1}{\tau}$ and $\gamma(\mu) - \gamma(0) = \mu\gamma'(0) + R$ with $\|R\|_2 \leq \frac{\mu^2}{2\tau}$. First, observe that for a geodesic curve, $\gamma'(t)$ is a parallel transport along γ [52]. Thus, $\|\gamma''\|_2 = \left\| \frac{d(\gamma')}{dt} \right\|_2 \leq \frac{\|\gamma'(0)\|_2}{\tau} = \frac{1}{\tau}$. Now, by the fundamental theorem of calculus, we have $\gamma(\mu) - \gamma(0) = \int_0^\mu \gamma'(t)\ dt$. Applying the fundamental theorem of calculus one more time on $\gamma'(t)$, we obtain

$$\gamma(\mu) - \gamma(0) = \mu\gamma'(0) + \int_0^\mu \int_0^t \gamma''(s)\ ds\ dt. \tag{49}$$

Then, applying $\|\gamma''(s)\|_2 \leq \frac{1}{\tau}$ and integrating completes the proof.

For the second part of the lemma, let $w(t)$ be the parameterization of the parallel transport of $u \in \mathcal{T}_p\mathcal{M}$ into $v \in \mathcal{T}_q\mathcal{M}$ such that $w(0) = u$ and $w(\mu) = v$, where $\mu := d_{\mathcal{M}}(p, q)$. Also, let $\gamma(t)$ be the unit speed parameterization of the geodesic curve between p and q as before. We want to show that $\langle u,\ v \rangle > 1 - \frac{\mu}{\tau}$. First, note that by definition of a parallel transport, we have for all t, $\|w(t)\|_2 = \|w(0)\|_2 = 1$. To obtain a lower bound on $\langle u, v \rangle$, we write

$$\langle u,\ v \rangle = \langle w(0),\ w(\mu) \rangle = \left\langle w(0),\ w(0) + \int_0^\mu w'(t)dt \right\rangle \geq 1 - \left| \left\langle w(0),\ \int_0^\mu w'(t)dt \right\rangle \right|,$$

where we invoked the fundamental theorem of calculus. Then by Cauchy-Schwarz, we have

$$\left|\left\langle w(0),\ \int_0^\mu w'(t)dt \right\rangle\right| \ \leq \ 1 \cdot \int_0^\mu \|w'(t)\|_2 \, dt.$$

Now by Lemma B.1.1, $\|w'(t)\|_2 \leq \frac{1}{\tau}$. Therefore, $\int_0^\mu \|w'(t)\|_2 \, dt \leq \frac{\mu}{\tau}$. The proof is complete after putting all the pieces together.

For the last part, again let $\gamma(t)$ be the unit speed parameterization of the geodesic curve connecting p and q. By applying the reverse triangle inequality on (49) and using the fact that $\|\gamma''(t)\|_2 \leq \frac{1}{\tau}$, we have

$$\|q - p\|_2 \geq \mu - \frac{\mu^2}{2\tau}.$$

The proof is complete after solving the above quadratic inequality for μ. □

B.2 Proof of Lemma 2.5.2

The proof of this lemma follows the arguments in the proof of [111, Proposition 10.1].

Proof. Let $\{b_1, \cdots, b_L\}$ the maximal $(\epsilon, d_{\mathcal{M}})$-packing of $\mathcal{M}$, defined as the largest set satisfying $b_l \in \mathcal{M}$ and $d_{\mathcal{M}}(b_l, b_m) > \epsilon$ for all $l, m \in [L]$ and $l \neq m$. By definition of the packing set, the geodesic balls of radius $\frac{\epsilon}{2}$ centered at the b_l's, i.e., $\left\{\mathcal{B}^{\mathcal{M}}_{d_{\mathcal{M}}}\left(b_l, \frac{\epsilon}{2}\right)\right\}_{l=1}^{L}$, do not intersect. Moreover, $\bigcup_{l=1,\cdots,L} \mathcal{B}^{\mathcal{M}}_{d_{\mathcal{M}}}\left(b_l, \frac{\epsilon}{2}\right) \subset \mathcal{M}$. Therefore,

$$\begin{aligned} &\text{vol}\left(\bigcup_{l=1,\cdots,L} \mathcal{B}^{\mathcal{M}}_{d_{\mathcal{M}}}\left(b_l, \frac{\epsilon}{2}\right)\right) = \sum_{l=1}^{L} \text{vol}\left(\mathcal{B}^{\mathcal{M}}_{d_{\mathcal{M}}}\left(b_l, \frac{\epsilon}{2}\right)\right) \leq \text{vol}(\mathcal{M}) =: V \\ \Rightarrow \quad & L \inf_{x \in \mathcal{M}} \text{vol}\left(\mathcal{B}^{\mathcal{M}}_{d_{\mathcal{M}}}\left(x, \frac{\epsilon}{2}\right)\right) \leq V. \end{aligned}$$

The proof of the first part of the lemma is complete after observing that the maximal $(\epsilon, d_{\mathcal{M}})$-packing of $\mathcal{M}$ is also a $(\epsilon, d_{\mathcal{M}})$-cover [111, Proposition 10.1]. Then, for any $x \in \mathcal{M}$, the geodesic regularity R of the manifold allows us to bound (see Definition 2.5.1)

$$\text{vol}\left(\mathcal{B}^{\mathcal{M}}_{d_{\mathcal{M}}}\left(x, \frac{\epsilon}{2}\right)\right) \geq \frac{1}{R^K} \text{vol}\left(\mathcal{B}^{K}_{2}\left(x, \frac{\epsilon}{2}\right)\right) = \frac{\sqrt{\pi}^K \epsilon^K}{(2R)^K \Gamma\left(\frac{K}{2} + 1\right)}.$$

With the following upper bound on the Gamma function [5]: $\Gamma\left(\frac{K}{2}+1\right) \leq \left(\frac{K}{2}+1\right)^{K/2}$, we have

$$\inf_{x \in \mathcal{M}} \operatorname{vol}\left(\mathcal{B}^{\mathcal{M}}_{d_{\mathcal{M}}}\left(x, \frac{\epsilon}{2}\right)\right) \geq \left(\frac{\sqrt{\pi}}{2R}\right)^{K} \left(\frac{\epsilon}{\sqrt{\frac{K}{2}+1}}\right)^{K}.$$

Putting this estimate with our covering number bound concludes the proof. □

APPENDIX C

BLOCK DIAGONAL MATRICES PROOFS

In this appendix, we prove the various theorems and lemmas appearing in Chapter 3.

C.1 Proof of Theorem 3.1.1 — Concentration of Measure for DBD Matrices

The proof of this theorem regarding the concentration of measure bounds for DBD matrices is a simple application of the Bernstein's inequality (Lemma A.1.5). Basically, we want to show that

$$\begin{aligned} & \mathbb{P}\left\{\left|\|\Phi x\|_2^2 - \|x\|_2^2\right| > \epsilon\|x\|_2^2\right\} \\ \leq \; & 2\exp\left(-C_1 \min\left\{\frac{C_2^2\epsilon^2}{\|\phi\|_{\psi_2}^4}\Gamma_2(x,\mathbf{M}), \frac{C_2\epsilon}{\|\phi\|_{\psi_2}^2}\Gamma_\infty(x,\mathbf{M})\right\}\right), \end{aligned}$$

First, let $y = \Phi x$, where we recall that

$$\Phi = \begin{pmatrix} \Phi_1 & & \\ & \ddots & \\ & & \Phi_J \end{pmatrix}.$$

For each matrix Φ_j, we let $[\Phi_j]_{i,n}$ denote the n-th entry of the i-th row of Φ_j. Further, we let $y_j(i)$ denote the i-th component of measurement vector y_j, and we let $x_j(n)$ denote the n-th entry of signal block x_j.

We begin by characterizing the point of concentration. One can write $y_j(i) = \sum_{n=1}^N [\Phi_j]_{i,n} x_j(n)$, and so it follows that $\mathbb{E}\left\{y_j^2(i)\right\} = \mathbb{E}\left\{\left(\sum_{n=1}^N [\Phi_j]_{i,n} x_j(n)\right)^2\right\}$. Since the $[\Phi_j]_{i,n}$ are zero-mean and independent, all cross product terms are equal to zero, and therefore we can write $\mathbb{E}\left\{y_j^2(i)\right\} = \mathbb{E}\left\{\sum_{n=1}^N [\Phi_j]_{i,n}^2 x_j^2(n)\right\} = \frac{1}{M_j}\|x_j\|_2^2$. Combining all of the measurements, we then have $\mathbb{E}\left\{\|y\|_2^2\right\} = \sum_{j=1}^J \sum_{i=1}^{M_j} \mathbb{E}\left\{y_j^2(i)\right\} = \sum_{j=1}^J \sum_{i=1}^{M_j} \frac{\|x_j\|_2^2}{M_j} = \sum_{j=1}^J \|x_j\|_2^2 = \|x\|_2^2$.

Now, we are interested in the probability that $|\|y\|_2^2 - \|x\|_2^2| > \epsilon\|x\|_2^2$. Since $\mathbb{E}\{\|y\|_2^2\} = \|x\|_2^2$, this is equivalent to the condition that $|\|y\|_2^2 - \mathbb{E}\{\|y\|_2^2\}| > \epsilon\mathbb{E}\{\|y\|_2^2\}$. For a given $j \in \{1, 2, \ldots, J\}$ and $i \in \{1, 2, \ldots, M_j\}$, all $\{[\Phi_j]_{i,n}\}_{n=1}^N$ are i.i.d. subgaussian random variables with subgaussian norms equal to $\left\|\frac{\phi}{\sqrt{M_j}}\right\|_{\psi_2} = \frac{\|\phi\|_{\psi_2}}{\sqrt{M_j}}$. From above, we know that $y_j(i)$ can be expressed as a linear combination of these random variables, with weights given by the entries of x_j. As with Gaussian random variables, linear combinations of i.i.d. subgaussian random variables are also subgaussian. In particular, from [136, Lemma 9] it follows that each $y_j(i)$ is a subgaussian random variable with subgaussian norm $\|y_j(i)\|_{\psi_2} \leq c_1 \frac{\|\phi\|_{\psi_2}}{\sqrt{M_j}} \|x_j\|_2$, where c_1 is an absolute constant.

In order to obtain a concentration bound for $\|y\|_2^2$, we shall apply the Bernstein inequality (Lemma A.1.5). First, let us define $\widetilde{y}_j(i) := \frac{y_j(i)}{\|y_j(i)\|_{\psi_2}}$ so that $\|\widetilde{y}_j(i)\|_{\psi_2} = 1$, and note that

$$\begin{aligned} & \mathbb{P}\left\{\left|\|y\|_2^2 - \mathbb{E}\left\{\|y\|_2^2\right\}\right| > \epsilon\|x\|_2^2\right\} \\ = \; & \mathbb{P}\left\{\left|\sum_j \sum_i \|y_j(i)\|_{\psi_2}^2 \left(\widetilde{y}_j^2(i) - \mathbb{E}\left\{\widetilde{y}_j^2(i)\right\}\right)\right| > \epsilon\|x\|_2^2\right\}. \end{aligned}$$

We apply Lemma A.1.5 to the subgaussian random variables $\widetilde{y}_j(i)$ (over all i, j) with weights $a_j(i) = \|y_j(i)\|_{\psi_2}^2$. Letting a denote a vector of length $\widetilde{M}$ containing these weights, we have that

$$\begin{aligned} \|a\|_2^2 &= \sum_j \sum_i a_j^2(i) = \sum_j \sum_i \|y_j(i)\|_{\psi_2}^4 \leq c_1^4 \|\phi\|_{\psi_2}^4 \sum_j \sum_i \frac{\|x_j\|_2^4}{M_j^2} \\ &= c_1^4 \|\phi\|_{\psi_2}^4 \sum_j \frac{\|x_j\|_2^4}{M_j} = c_1^4 \|\phi\|_{\psi_2}^4 \|\mathbf{M}^{-1/2}\gamma\|_2^2, \end{aligned}$$

and

$$\|a\|_\infty = \max_{i,j} a_j(i) = \max_{i,j} \|y_j(i)\|_{\psi_2}^2 \leq c_1^2 \|\phi\|_{\psi_2}^2 \max_j \frac{\|x_j\|_2^2}{M_j} = c_1^2 \|\phi\|_{\psi_2}^2 \|\mathbf{M}^{-1}\gamma\|_\infty.$$

Further note that $\|x\|_2^2 = \|\gamma\|_1$ and $\|x\|_2^4 = \|\gamma\|_1^2$. We complete the proof by substituting these quantities into Lemma A.1.5 with $T = 1$ and $t = \epsilon\|x\|_2^2$ and by taking $C_2 = \frac{1}{4c_1^2}$.

C.2 Proof of Theorem 3.1.2 — Concentration of Measure for RBD Matrices

Again, the proof of this theorem follows from Bernstein's inequality (Lemma A.1.5) after some manipulation. We want to show that

$$\begin{aligned} & \mathbb{P}\left\{\left|\|\Phi x\|_2^2 - \|x\|_2^2\right| > \epsilon\|x\|_2^2\right\} \\ \leq \; & 2\exp\left(-C_1 \min\left\{C_3^2\epsilon^2\Lambda_2(x,M),\; C_3\epsilon\Lambda_\infty(x,M)\right\}\right), \end{aligned}$$

For this, let $y = \Phi x$ with

$$\Phi = \begin{pmatrix} \widetilde{\Phi} & & \\ & \ddots & \\ & & \widetilde{\Phi} \end{pmatrix},$$

and let the blocks on the diagonal of Φ be $\widetilde{\Phi}^T = [\phi_1 \mid \phi_2 \mid \cdots \mid \phi_M]$ with each $\phi_i \in \mathbb{R}^N$ being an i.i.d. Gaussian random vector. Then,

$$\begin{aligned} \|y\|_2^2 &= x^T\Phi^T\Phi x = \sum_{j=1}^{J} x_j^T\widetilde{\Phi}^T\widetilde{\Phi}x_j = \sum_{j=1}^{J} x_j^T\left(\sum_{m=1}^{M}\phi_i\phi_i^T\right)x_j \\ &= \sum_{i=1}^{M}\phi_m^T\left(\sum_{j=1}^{J}x_jx_j^T\right)\phi_m = \sum_{m=1}^{M}\phi_m^TA\phi_m. \end{aligned}$$

Because A is symmetric, it has an eigen-decomposition $A = V^TDV$, where D is a diagonal matrix of its eigenvalues $\{\lambda_n\}_{n=1}^N$ and V is an orthogonal matrix of eigenvectors. Then, for any $m = 1, \cdots, M$, we have

$$\phi_m^TA\phi_m = (V\phi_m)^TD(V\phi_m) = \sum_{n=1}^{N}\lambda_n\phi_m(n)^2,$$

where $\widetilde{\phi}_m = V\phi_m$ with $\widetilde{\phi}_m = [\phi_m(1),\; \cdots,\; \phi_m(N)]^T$. Since V is an orthogonal matrix, $\{\widetilde{\phi}_m(n)\}_{n=1}^N$ are i.i.d. Gaussian random variables with zero-mean and variance $\sigma^2 = \frac{1}{M}$. First, notice that

$$\mathbb{E}\left\{\phi_m^TA\phi_m\right\} = \mathbb{E}\left\{\sum_{n=1}^{N}\lambda_n\phi_m(n)^2\right\} = \frac{1}{M}\sum_{n=1}^{N}\lambda_n = \frac{\|X\|_F^2}{M} = \frac{\|x\|_2^2}{M}.$$

Thus,

$$\begin{aligned}
\|y\|_2^2 - \|x\|_2^2 &= \sum_{m=1}^{M} \phi_m^T A \phi_m - \|x\|_2^2 \\
&= \sum_{m=1}^{M} \left(\phi_m^T A \phi_m - \mathbb{E}\left\{ \phi_m^T A \phi_m \right\} \right) \\
&= \sum_{m=1}^{M} \left(\sum_{n=1}^{N} \lambda_n \phi_m(n)^2 + \mathbb{E}\left\{ \sum_{n=1}^{N} \lambda_n \phi_m(n)^2 \right\} \right) \\
&= \sum_{m=1}^{M} \sum_{n=1}^{N} \lambda_n \left(\phi_m(n)^2 - \mathbb{E}\left\{ \phi_m(n)^2 \right\} \right).
\end{aligned}$$

We can then apply Bernstein's inequality (Lemma A.1.5) with the subgaussian sequence $\{\phi_m(n)\}_{m,n}$. For the terms appearing in the lemma, we see that $T := \max_{m,n} \|\phi_m(n)\|_{\psi_2} \leq \frac{C}{M}$ for some constant C, $\|a\|_2^2 = M\|\lambda\|_2^2$, $\|a\|_\infty = \|\lambda\|_\infty$. Finally, substituting $t = \epsilon\|x\|_2^2 = \epsilon\|\lambda\|_1$ into the tail bound completes our proof.

C.3 Proof of Theorem 3.1.6

In this section, we want to show that if $x \in \mathbb{R}^{\widetilde{N}}$ has frequency support $\Omega \subset [\widetilde{N}]$ of size $S \leq N$ generated uniformly at random, then with probability at least $1 - O(J(\log(\widetilde{N}))^{1/2}(\widetilde{N})^{-\beta})$,

$$\Gamma_2(x, \mathbf{M}) \geq \widetilde{M} \cdot \min \left\{ \frac{0.0779}{(\beta+1)\log(\widetilde{N})}, \frac{1}{\left(\sqrt{6(\beta+1)\log \widetilde{N}} + \frac{(\log \widetilde{N})^2}{N} \right)^2} \right\},$$

where $\widetilde{N} > 512$ and $\mathbf{M} = \mathrm{diag}\{M, M, \ldots, M\}$. This result follows from an application of the following:

Theorem C.3.1. *[28, Theorem 3.1] Let $x \in \mathbb{C}^{\widetilde{N}}$ and $\beta > 1$. Suppose $\widetilde{N} > 512$ and choose q, N_T, and N_Ω such that*

$$N_T + N_\Omega \leq \frac{0.5583\widetilde{N}/q}{\sqrt{(\beta+1)\log(\widetilde{N})}} \text{ and } N_T + N_\Omega \leq \frac{\sqrt{2/3}\widetilde{N}\left(\frac{1}{q} - \frac{(\log \widetilde{N})^2}{\widetilde{N}}\right)}{\sqrt{(\beta+1)\log(\widetilde{N})}}. \tag{50}$$

Fix a subset T of the time domain with $|T| = N_T$. Let Ω be a subset of size N_Ω of the frequency domain generated uniformly at random. Then with probability at least $1 - O((\log(\widetilde{N}))^{1/2}\widetilde{N}^{-\beta})$, every signal x supported on Ω in the frequency domain has most of its energy in the time domain outside of T. In particular, $\|x_T\|_2^2 \leq \frac{\|x\|_2^2}{q}$, where x_T denotes the restriction of x to the support T.

Proof. First, observe that $\|\gamma\|_1^2 = \|x\|_2^4$ and $\|\gamma\|_2^2 = \sum_{k=1}^J \|x_k\|_2^4$. Next, apply Theorem C.3.1 with $N_\Omega = S$ and $N_T = N = \widetilde{N}/J$, being careful to select a value for q such that (50) is satisfied. In particular, we require

$$\frac{1}{q} \geq \frac{(N+S)\sqrt{(\beta+1)\log\widetilde{N}}}{0.5583\widetilde{N}} \quad \text{and} \quad \frac{1}{q} \geq \frac{\frac{(N+S)}{\sqrt{2/3}}\sqrt{(\beta+1)\log\widetilde{N}} + (\log\widetilde{N})^2}{\widetilde{N}}.$$

This is satisfied if we choose

$$q \leq \min\left\{ \frac{0.5583\widetilde{N}}{(N+S)\sqrt{(\beta+1)\log\widetilde{N}}}, \frac{\widetilde{N}}{\frac{(N+S)}{\sqrt{2/3}}\sqrt{(\beta+1)\log\widetilde{N}} + (\log\widetilde{N})^2} \right\}. \tag{51}$$

Choosing any q satisfying (51), we have that $\|x_k\|_2^2 \leq \frac{\|x\|_2^2}{q}$ for each $k \in [J]$ with failure probability at most $O((\log(\widetilde{N}))^{1/2}(\widetilde{N})^{-\beta})$, implying that each block individually is favorable. Taking a union bound for all k to cover each block, we have that with total failure probability at most $O(J(\log(\widetilde{N}))^{1/2}(\widetilde{N})^{-\beta})$, $\|\gamma\|_2^2 = \sum_{k=1}^J \|x_k\|_2^4 \leq \frac{J\|x\|_2^4}{q^2}$. Thus with this same failure probability, $\frac{\Gamma_2}{M} = \frac{\|\gamma\|_1^2}{\|\gamma\|_2^2} \geq \frac{q^2}{J}$. Combining with (51) and using the fact that $S < N$, we thus have:

$$\begin{aligned}
\frac{\Gamma_2}{M} &\geq \min\left\{ \frac{0.5583^2 N^2 J}{(N+S)^2(\beta+1)\log\widetilde{N}}, \frac{N^2 J}{\left(\frac{(N+S)}{\sqrt{2/3}}\sqrt{(\beta+1)\log\widetilde{N}} + (\log\widetilde{N})^2\right)^2} \right\} \\
&\geq \min\left\{ \frac{(0.5583^2/2^2)J}{(\beta+1)\log(\widetilde{N})}, \frac{J}{\left(\frac{2}{\sqrt{2/3}}\sqrt{(\beta+1)\log\widetilde{N}} + \frac{(\log\widetilde{N})^2}{N}\right)^2} \right\}.
\end{aligned}$$

□

C.4 Proof of Lemma 3.1.2

In this appendix, we want to show that if $x \in \mathbb{R}^{\widetilde{N}}$ is a vector whose entries are i.i.d. subgaussian random variables with mean 0, variance σ^2, and subgaussian norm $\|\phi\|_{\psi_2}$, then with probability at least $1 - 2\exp\left(-\frac{1}{2}\frac{C_1 C_2^2 N\epsilon^2}{\|\phi\|_{\psi_2}^4}\right)$,

$$\Gamma_2(x, \mathbf{M}) \geq \Lambda_2(x, M) \geq M + M\left(\frac{1-\epsilon}{1+\epsilon}\right)^2 (J-1),$$

where $\epsilon \leq \frac{\|\phi\|_{\psi_2}^2}{C_2}$, $J \leq \frac{C_1 C_2^2 N\epsilon^2}{2\|\phi\|_{\psi_2}^4 \log(12/\epsilon)}$, and $\mathbf{M} = \text{diag}\{M, M, \ldots, M\}$.

Proof. Let X be the $J \times N$ matrix as defined in (15). Without loss of generality, we suppose the nonzero eigenvalues $\{\lambda_i\}_{i=1}^{\min(J,N)}$ of $X^T X$ are sorted in order of decreasing magnitude, and we let $\lambda_{\max} := \lambda_1$ and $\lambda_{\min} := \lambda_{\min(J,N)}$. We can lower bound Λ_2 in terms of these extremal eigenvalues by writing

$$\begin{aligned}
\Lambda_2 &= \frac{M\|\lambda\|_1^2}{\|\lambda\|_2^2} = M\frac{\sum_i \lambda_i^2 + \sum_i \sum_{j\neq i} \lambda_i\lambda_j}{\sum_i \lambda_i^2} \\
&\geq M + M\frac{\lambda_{\min}}{\lambda_{\max}}\frac{\sum_i \sum_{j\neq i} \lambda_i}{\sum_i \lambda_i} = M + M\frac{\lambda_{\min}}{\lambda_{\max}}(J-1).
\end{aligned}$$

Assume that $\epsilon \leq \frac{\|\phi\|_{\psi_2}^2}{C_2}$, and let us define the following events:

$$\begin{aligned}
A &= \left\{N\sigma^2(1-\epsilon)^2 \leq \frac{\|X^T z\|_2^2}{\|z\|_2^2} \leq N\sigma^2(1+\epsilon)^2,\ \forall z \in \mathbb{R}^J\right\}, \\
B &= \left\{\lambda_{\max} \leq N\sigma^2(1+\epsilon)^2\right\} \bigcap \left\{\lambda_{\min} \geq N\sigma^2(1-\epsilon)^2\right\}, \\
C &= \left\{\frac{\lambda_{\min}}{\lambda_{\max}} \geq \left(\frac{1-\epsilon}{1+\epsilon}\right)^2\right\}, \\
D &= \left\{\Lambda_2 \geq M + M\left(\frac{1-\epsilon}{1+\epsilon}\right)^2 (J-1)\right\}.
\end{aligned}$$

These events satisfy $A = B \subseteq C \subseteq D$, where the last relation follows from (52). It follows that $\mathbb{P}\{D^c\} \leq \mathbb{P}\{A^c\}$, where A^c represents the complement of event A. Because X^T is populated with i.i.d. subgaussian random variables, it follows as a corollary of Theorem 3.1.1 (by setting $M \leftarrow N$ and $J \leftarrow 1$ in the context of that theorem) that for any $z \in^J$ and $\epsilon \leq \frac{\|\phi\|_{\psi_2}^2}{C_2}$, $\mathbb{P}\left\{\left|\|X^T z\|_2^2 - N\sigma^2\|z\|_2^2\right| > \epsilon N\sigma^2\|z\|_2^2\right\} \leq 2\exp\left(-\frac{C_1 C_2^2 N\epsilon^2}{\|\phi\|_{\psi_2}^4}\right)$.

Thus, for an upper bound for $P(A^c)$, we can follow the straightforward arguments in [13, Lemma 5.1] and conclude that $\mathbb{P}\{A^c\} \leq 2\left(\frac{12}{\epsilon}\right)^J \exp\left(-\frac{C_1 C_2^2 N\epsilon^2}{\|\phi\|_{\psi_2}^4}\right)$. Thus by choosing $J \leq \frac{C_1 C_2^2 N\epsilon^2}{2\|\phi\|_{\psi_2}^4 \log(12/\epsilon)}$, we see that $\mathbb{P}\{D^c\} \leq 2\exp\left(-\frac{1}{2}\frac{C_1 C_2^2 N\epsilon^2}{\|\phi\|_{\psi_2}^4}\right)$. Finally, the fact that $\Gamma_2 \geq \Lambda_2$ follows from (18). □

C.5 Proof of Lemma 3.2.1

In this section, we want to show that if Ψ is a random orthonormal basis as described in Section 3.2.1, then, for a fixed $t \gtrsim 1$ and $\widetilde{N} \gtrsim t^2 \log \widetilde{N}$, we have

$$\mathbb{P}\left\{\mu(\Psi) > t\sqrt{\log \widetilde{N}}\right\} \lesssim \widetilde{N}^{-t}.$$

Let $g \in \mathbb{R}^{\widetilde{N}}$ denote a vector whose entries are i.i.d. zero-mean Gaussian random variables with unit variance. Note that ψ_1 is drawn from the uniform distribution on the unit sphere in $\mathbb{R}^{\widetilde{N}}$. Therefore the entries of ψ_1 has the same distribution as those in $\frac{g}{\|g\|_2}$. Since the distribution of Ψ remains unchanged under permutation of its rows, every column of Ψ has the same (marginal) distribution as $\frac{g}{\|g\|_2}$. For any $t > 0$, this,

along with the union bound, allows us to write

$$\begin{aligned}
\mathbb{P}\left\{\mu\left(\Psi\right) > t\sqrt{\log\widetilde{N}}\right\} &= \mathbb{P}\left\{\max_{\widetilde{n}\in[\widetilde{N}]}\|\psi_{\widetilde{n}}\|_\infty > t\sqrt{\log\widetilde{N}}/\sqrt{\widetilde{N}}\right\} \\
&\leq \widetilde{N}\cdot\max_{\widetilde{n}\in[\widetilde{N}]}\mathbb{P}\left\{\|\psi_{\widetilde{n}}\|_\infty > t\sqrt{\log\widetilde{N}}/\sqrt{\widetilde{N}}\right\} \\
&= \widetilde{N}\cdot\mathbb{P}\left\{\max_{\widetilde{n}\in[\widetilde{N}]}\frac{|g(\widetilde{n})|}{\|g\|_2} > t\sqrt{\frac{\log\widetilde{N}}{\widetilde{N}}}\right\} \\
&\leq \widetilde{N}^2\cdot\mathbb{P}\left\{\frac{|g(1)|}{\|g\|_2} > t\sqrt{\frac{\log\widetilde{N}}{\widetilde{N}}}\right\} \\
&= \widetilde{N}^2\cdot\mathbb{P}\left\{\frac{|g(1)|}{\|g\|_2} > \frac{t/2\cdot\sqrt{\log\widetilde{N}}}{(1-1/2)\sqrt{\widetilde{N}}}\right\} \\
&\leq \widetilde{N}^2\cdot\mathbb{P}\left\{|g(1)| > t/2\cdot\sqrt{\log\widetilde{N}}\right\} + \widetilde{N}^2\cdot\mathbb{P}\left\{\|g\|_2 < (1-1/2)\sqrt{\widetilde{N}}\right\} \\
&\leq \widetilde{N}^2 e^{-\frac{C}{4\|g(1)\|_{\psi_2}^2}t^2\log\widetilde{N}} + \widetilde{N}^2 e^{-C\min\left(\frac{1}{4\|g(1)\|_{\psi_2}^4},\frac{1}{2\|g(1)\|_{\psi_2}^2}\right)\widetilde{N}} \\
&\leq \widetilde{N}^2 e^{-\frac{C}{4\|g(1)\|_{\psi_2}^2}t^2\log\widetilde{N}} + \widetilde{N}^2 e^{-\frac{C}{4\|g(1)\|_{\psi_2}^4}\widetilde{N}},
\end{aligned}$$

where we used the tail bound of Gaussian random variables and Bernstein's inequality (Lemma A.1.5) to bound the failure probability. The last line holds because $\|g(1)\|_{\psi_2} = \sqrt{2/\pi}$.[1] If we take $\widetilde{N} \geq t^2\log\widetilde{N}$, we obtain that

$$\mathbb{P}\left\{\mu\left(\Psi\right) > t\sqrt{\log\widetilde{N}}\right\} \leq \widetilde{N}^{2-\frac{C}{4\|g(1)\|_{\psi_2}^2}t^2} + \widetilde{N}^{2-\frac{C}{4\|g(1)\|_{\psi_2}^4}t^2}.$$

We arrive at the advocated result with an appropriate choice of $t \gtrsim 1$, i.e.,

$$\mathbb{P}\left\{\mu\left(\Psi\right) > t\sqrt{\log\widetilde{N}}\right\} \leq \widetilde{N}^{-t} + \widetilde{N}^{-t} = 2\widetilde{N}^{-t}.$$

[1] This is easily verified using the moments of Gaussian distribution.

C.6 Proof of Lemma 3.2.2

In this section, we want to show that for the random orthonormal basis Ψ constructed in Section 3.2.1, we have

$$\mathbb{P}\left\{\gamma\left(\Psi\right) \gtrsim 1+\sqrt{\frac{J}{N}}+t\right\} \lesssim \widetilde{N}^{-1},$$

if $J \leq N$ and $N \gtrsim t^{-2}$ for a fixed $t \leq 1$.

As pointed out in the proof of Lemma 3.2.1, the columns of Ψ are dependent but identically distributed as $\frac{g}{\|g\|_2}$, where g was defined there. Now, for every t, we can write

$$\begin{aligned}
\mathbb{P}\left\{\gamma\left(\Psi\right) \gtrsim 1+\sqrt{\frac{J}{N}}+t\right\} &= \mathbb{P}\left\{\max_{\widetilde{n}\in[\widetilde{N}]} \|X_R(\Psi, e_{\widetilde{n}})\|_2 \gtrsim \frac{1}{\sqrt{N}}+\frac{1+t}{\sqrt{J}}\right\} \\
&\leq \widetilde{N} \cdot \max_{\widetilde{n}\in[\widetilde{N}]} \mathbb{P}\left\{\|X_R(\Psi, e_{\widetilde{n}})\|_2 \gtrsim \frac{1}{\sqrt{N}}+\frac{1+t}{\sqrt{J}}\right\} \\
&= \widetilde{N} \cdot \mathbb{P}\left\{\|X_R(\Psi, e_1)\|_2 \gtrsim \frac{1}{\sqrt{N}}+\frac{1+t}{\sqrt{J}}\right\}. \qquad (52)
\end{aligned}$$

The second line uses the union bound and the last line holds due to the identical distribution of the columns of Ψ. It remains to find an upper bound for the probability in the last line above. Recall that ψ_1 has the same distribution as $\frac{g}{\|g\|_2}$, and thus $X_R(\Psi, e_1)$ has the same distribution as $G/\|G\|_F$, where $G \in \mathbb{C}^{N\times J}$ is formed by reshaping g. Therefore, $\|X_R(\Psi, e_1)\|_2$ has the same distribution as $\|G\|_2 / \|G\|_F$. For fixed $t \leq 1$, the following convenient inequality holds:

$$\frac{1}{\sqrt{N}}+\frac{1+t}{\sqrt{J}} \gtrsim \frac{(1+t/3)\sqrt{N}+\sqrt{J}}{(1-t/3)\sqrt{\widetilde{N}}}. \qquad (53)$$

Now, we can write that

$$\begin{aligned}
&\mathbb{P}\left\{ \|X_R(\Psi, e_1)\|_2 \gtrsim \frac{1}{\sqrt{N}} + \frac{1+t}{\sqrt{J}} \right\} \\
&= \mathbb{P}\left\{ \frac{\|G\|_2}{\|G\|_F} \gtrsim \frac{1}{\sqrt{N}} + \frac{1+t}{\sqrt{J}} \right\} \\
&\leq \mathbb{P}\left\{ \frac{\|G\|_2}{\|G\|_F} \gtrsim \frac{\left(1+\frac{t}{3}\right)\sqrt{N} + \sqrt{J}}{\left(1-\frac{t}{3}\right)\sqrt{\widetilde{N}}} \right\} \\
&\leq \mathbb{P}\left\{ \|G\|_2 \gtrsim \left(1+\frac{t}{3}\right)\sqrt{N} + \sqrt{J} \right\} + \mathbb{P}\left\{ \|G\|_F \lesssim \left(1-\frac{t}{3}\right)\sqrt{\widetilde{N}} \right\} \\
&\leq e^{-\frac{1}{18}t^2 N} + e^{-C\min\left(\frac{t^2}{9\|g(1)\|_{\psi_2}^4}, \frac{t}{3\|g(1)\|_{\psi_2}^2}\right)\widetilde{N}} \\
&\leq e^{-\frac{1}{18}t^2 N} + e^{-\frac{C}{9\|g(1)\|_{\psi_2}^4} t^2 \widetilde{N}} \\
&\leq 2e^{-Ct^2 N},
\end{aligned}$$

where the second line uses (53). The fourth line uses the properties of a Gaussian random matrix (see [136]) and Bernstein's inequality (Lemma A.1.5). The second to last line follows because $\|g(1)\|_{\psi_2} = \sqrt{2/\pi}$ and $t \leq 1$. The above upper bound in combination with (52) leads us to the conclusion that

$$\mathbb{P}\left\{ \gamma(\Psi) \gtrsim 1 + \sqrt{\frac{J}{N}} + t \right\} \leq 2\widetilde{N} e^{-Ct^2 N}.$$

We complete the proof of Lemma 3.2.2 by taking $N \geq 3C^{-1}t^{-2}\log\widetilde{N}$.

C.7 Proofs of Theorems 3.2.1 and 3.2.2

In this section, we show that for a $\widetilde{M} \times \widetilde{N}$ DBD or RBD matrix Φ and a basis Ψ, $\Phi\Psi$ satisfies the RIP-(S, δ) with high probability whenever $J \gtrsim 1$ and

$$\widetilde{M} \gtrsim_\tau \frac{S}{\delta^2} \kappa(\Psi)^2 \log^2 S \log^2 \widetilde{N}.$$

Here, $\kappa(\Psi) := \widetilde{\mu}(\Psi)$ for a DBD matrix or $\kappa(\Psi) := \gamma(\Psi)$ for a RBD matrix.

C.7.1 Preliminaries

First, leveraging the fact that $\|\alpha\|_2 = 1$ for $\alpha \in \Sigma_S$, we have

$$\mathbb{E}\left\{ \|\Phi \cdot x(\alpha)\|_2^2 \right\} = \alpha^H \Psi^H \mathbb{E}\left\{ \Phi^H \Phi \right\} \Psi\alpha = \alpha^H \alpha = 1.$$

for Φ representing either a DBD or RBD matrix. Then, define

$$\delta_S \;:=\; \sup_{\alpha\in\Sigma_S} \left| \|\Phi\cdot x(\alpha)\|_2^2 - 1 \right|,$$

where we observe that δ_S is the supremum of a random process $|\|\Phi\cdot x(\alpha)\|_2^2-1|$ over the index set Σ_S. Given a predetermined conditioning $\delta<1$ and under the conditions of Theorems 3.2.1 and 3.2.2, our objective is to show that $\delta_S\leq\delta$ for both DBD and RBD matrices. To achieve this goal, we require the following result due to Krahmer et al.:

Theorem C.7.1. *[83, Theorem 3.1] Let $\mathcal{A}\subset\mathbb{C}^{\widetilde{M}\times\widetilde{N}}$ be a set of matrices, and let ξ be a random vector whose entries are i.i.d., zero-mean, unit-variance random variables with subgaussian norm τ. Set*

$$d_F(\mathcal{A}) := \sup_{A\in\mathcal{A}} \|A\|_F, \quad d_2(\mathcal{A}) := \sup_{A\in\mathcal{A}} \|A\|_2,$$

and

$$\begin{aligned} E_1 &:= \gamma_2\left(\mathcal{A},\|\cdot\|_2\right)\left(\gamma_2\left(\mathcal{A},\|\cdot\|_2\right)+d_F(\mathcal{A})\right)+d_F(\mathcal{A})d_2(\mathcal{A}),\\ E_2 &:= d_2(\mathcal{A})\left(\gamma_2\left(\mathcal{A},\|\cdot\|_2\right)+d_F(\mathcal{A})\right),\\ E_3 &:= d_2^2(\mathcal{A}). \end{aligned}$$

Then, for $\epsilon>0$, it holds that

$$\log\mathbb{P}\left\{\sup_{A\in\mathcal{A}}\left|\|A\xi\|_2^2-\mathbb{E}\left\{\|A\xi\|_2^2\right\}\right|\gtrsim_\tau E_1+\epsilon\right\}\lesssim_\tau -\min\left(\frac{\epsilon^2}{E_2^2},\frac{\epsilon}{E_3}\right).$$

Without going into the details, we note that the γ_2-functional of $\mathcal{A}$, $\gamma_2\left(\mathcal{A},\|\cdot\|_2\right)$, is a geometrical property of $\mathcal{A}$ that is widely used in the context of probability in Banach spaces [88, 127]. In particular, the following lemma gives an estimate of this quantity.

Lemma C.7.1. [83, 127] *For $\mathcal{A}$ defined in Theorem C.7.1, we have*

$$\gamma_2\left(\mathcal{A},\|\cdot\|_2\right)\lesssim\int_0^\infty \log^{\frac{1}{2}}\left(\left|\mathcal{C}\left(\mathcal{A},\|\cdot\|_2,\eta\right)\right|\right)\,d\eta. \tag{54}$$

Clearly, we need to express the problem of bounding δ_S in a form that is amenable to the setting of Theorem C.7.1. First, for DBD matrices Φ, let us define $X_{D,j} \in \mathbb{C}^{M \times MN}$, $j \in [J]$, as

$$X_{D,j}(\alpha) = X_{D,j}(\alpha, \Psi) := \begin{bmatrix} x_j^H(\alpha) & & & \\ & x_j^H(\alpha) & & \\ & & \ddots & \\ & & & x_j^H(\alpha) \end{bmatrix}. \tag{55}$$

Then, we see that

$$\begin{aligned} \|\Phi \cdot x(\alpha)\|_2^2 &= \sum_{j \in [J]} \|\Phi_j \cdot x_j(\alpha)\|_2^2 \\ &= \sum_{j \in [J]} \|X_{D,j}(\alpha) \cdot \mathrm{vec}\left(\Phi_j^T\right)\|_2^2 \\ &\overset{\text{dist.}}{=} \sum_{j \in [J]} \|\frac{1}{\sqrt{M}} X_{D,j}(\alpha) \cdot \xi_j\|_2^2 \\ &\overset{\text{dist.}}{=} \|A_D(\alpha) \cdot \xi\|^2, \end{aligned}$$

where $\mathrm{vec}\left(\Phi_j^H\right)$ converts the $N \times M$ matrix Φ_j^T into a length-NM vector by stacking its columns. The linear map $A_D : \Sigma_S \to \mathbb{C}^{\widetilde{M} \times JMN}$ is defined as

$$A_D(\alpha) = A_D(\alpha, \Psi) := \frac{1}{\sqrt{M}} \begin{bmatrix} X_{D,1}(\alpha) & & & \\ & X_{D,2}(\alpha) & & \\ & & \ddots & \\ & & & X_{D,J}(\alpha) \end{bmatrix},$$

and $\xi = \left[\xi_1^T, \cdots, \xi_J^T\right]^T \in \mathbb{R}^{JMN}$ (with $\xi_j \in \mathbb{R}^{MN}$) is composed of i.i.d. zero-mean, unit-variance subgaussian random variables with subgaussian norm τ. The index set $\mathcal{A}$ in Theorem C.7.1 for DBD matrices is $\mathcal{A}_D := \{A_D(\alpha) : \alpha \in \Sigma_S\}$. Thus, we have expressed the DBD RIP problem in the setting of Theorem C.7.1.

Next, for RBD matrices, we observe that

$$\begin{aligned}\|\Phi \cdot x(\alpha)\|_2^2 &= \sum_{j\in[J]} \|\Phi_1 \cdot x_j(\alpha)\|_2^2 \\ &= \|\Phi_1 \cdot X_R(\alpha)\|_F^2 \\ &= \|X_R^H(\alpha) \cdot \Phi_1^T\|_F^2 \\ &\overset{\text{dist.}}{=} \sum_{m\in[M]} \left\| \frac{1}{\sqrt{M}} X_R^H(\alpha) \cdot \xi_m \right\|_2^2 \\ &\overset{\text{dist.}}{=} \|A_R(\alpha) \cdot \xi\|_2^2, \end{aligned} \tag{56}$$

where we used the fact that $\Phi_1 = \Phi_2 = \cdots = \Phi_J$ and the linear map $A_R : \Sigma_S \to \mathbb{C}^{\widetilde{M}\times MN}$ is defined as

$$A_R(\alpha) = A_R(\alpha, \Psi) := \frac{1}{\sqrt{M}} \begin{bmatrix} X_R^H(\alpha) & & & \\ & X_R^H(\alpha) & & \\ & & \ddots & \\ & & & X_R^H(\alpha) \end{bmatrix}, \tag{57}$$

and $\xi = \left[\xi_1^T, \cdots, \xi_J^T\right]^T \in \mathbb{R}^{JMN}$ (with $\xi_j \in \mathbb{R}^{MN}$) is composed of i.i.d. zero-mean, unit-variance subgaussian random variables with subgaussian norm τ. The index set $\mathcal{A}$ in Theorem C.7.1 for RBD matrices is now $\mathcal{A}_R := \{A_R(\alpha) : \alpha \in \Sigma_S\}$. Thus, we have expressed the RBD RIP problem in the setting of Theorem C.7.1. In the next two subsections, we will estimate the various quantities involved in Theorem C.7.1 for both the DBD and RBD RIP problems.

C.7.2 Calculating $d_2(\mathcal{A}_D)$, $d_F(\mathcal{A}_D)$, and $\gamma_2(\mathcal{A}_D, \|\cdot\|_2)$

To begin, let us define the A_D-norm on a vector $\alpha \in \mathbb{C}^{\widetilde{N}}$ as

$$\|\alpha\|_{A_D} := \|A_D(\alpha)\|_2. \tag{58}$$

The A_D-norm can be bounded by the ℓ_1-norm as follows:

Lemma C.7.2. *For every $\alpha \in \mathbb{C}^{\widetilde{N}}$, it holds that*

$$\|\alpha\|_{A_D} \leq \frac{\widetilde{\mu}}{\sqrt{\widetilde{M}}} \cdot \|\alpha\|_1.$$

Proof. Let $\psi_{j,n}$ for $j \in [J]$ and $n \in [N]$ denote the $(j-1)N+n$ row of Ψ. We then have that

$$\begin{aligned}\|\alpha\|_{A_D} = \|A_D(\alpha)\|_2 = \|A_D(\alpha)A_D^H(\alpha)\|_2^{\frac{1}{2}} &= \frac{1}{\sqrt{M}} \max_{j\in[J]} \|x_j\|_2 \\ &= \frac{1}{\sqrt{M}} \max_{j\in[J]} \|\Psi_j \alpha\|_2 \leq \sqrt{\frac{N}{M}} \max_{j\in[J],n\in[N]} |\langle \psi_{j,n}, \alpha\rangle| \\ &\leq \sqrt{\frac{N}{M}} \max_{j\in[J],n\in[N]} \|\psi_{j,n}\|_\infty \|\alpha\|_1 = \frac{\mu}{\sqrt{\widetilde{M}}} \cdot \|\alpha\|_1,\end{aligned}$$

where we used the Hölder inequality and the definition of $\mu = \mu(\Psi)$. On the other hand, one may also write that

$$\|\alpha\|_{A_D} = \frac{1}{\sqrt{M}} \max_{j\in[J]} \|x_j\|_2 \leq \frac{1}{\sqrt{M}} \|x\|_2 = \frac{1}{\sqrt{M}} \|\Psi\alpha\|_2 = \frac{1}{\sqrt{M}} \|\alpha\|_2 \leq \frac{1}{\sqrt{M}} \|\alpha\|_1,$$

where we used the fact that Ψ is an basis. Thus, by using the definition of $\widetilde{\mu}$, we arrive at

$$\|\alpha\|_{A_D} \leq \frac{1}{\sqrt{\widetilde{M}}} \min\left(\mu, \sqrt{J}\right) \|\alpha\|_1 = \frac{\widetilde{\mu}}{\sqrt{\widetilde{M}}} \cdot \|\alpha\|_1.$$

□

With the A_D-norm characterized, let us begin calculating the quantities involved in Theorem C.7.1. First, we have

$$d_F(\mathcal{A}_D) = \sup_{A_D(\alpha)\in\mathcal{A}_D} \|A_D(\alpha)\|_F = \sup_{\alpha\in\Sigma_S} \|x(\alpha)\|_2 = \sup_{\alpha\in\Sigma_S} \|\Psi\alpha\|_2 = \sup_{\alpha\in\Sigma_S} \|\alpha\|_2 = 1. \quad (59)$$

We used again the fact that Ψ is a basis. Second, we have that

$$d_2(\mathcal{A}_D) = \sup_{A_D(\alpha)\in\mathcal{A}_D} \|A_D(\alpha)\|_2 = \sup_{\alpha\in\Sigma_S} \|\alpha\|_{A_D} \leq \frac{\widetilde{\mu}}{\sqrt{\widetilde{M}}} \sup_{\alpha\in\Sigma_S} \|\alpha\|_1 \leq \widetilde{\mu}\sqrt{\frac{S}{\widetilde{M}}}, \quad (60)$$

where we used Lemma C.7.2 and the fact that $\alpha \in \Sigma_S$ for bounding its ℓ_1-norm. It remains to bound $\gamma_2(\mathcal{A}_D, \|\cdot\|_2)$. According to Lemma C.7.1, we have that

$$\begin{aligned}\gamma_2(\mathcal{A}_D, \|\cdot\|_2) &\leq \int_0^\infty \log^{\frac{1}{2}}\left(\left|\mathcal{C}\left(\mathcal{A}_D, \|\cdot\|_2, \eta\right)\right|\right) \, d\eta \\ &= \int_0^\infty \log^{\frac{1}{2}}\left(\left|\mathcal{C}\left(\Sigma_S, \|\cdot\|_{A_D}, \eta\right)\right|\right) \, d\eta,\end{aligned}$$

where we used the equivalence between $\mathcal{A}_D$ with metric $\|\cdot\|_2$ and Σ_S with metric $\|\cdot\|_{A_D}$. Continuing on our bounds, we have

$$\begin{aligned}\gamma_2(\mathcal{A}_D, \|\cdot\|_2) &\leq \int_0^\infty \log^{\frac{1}{2}}\left(\left|\mathcal{C}\left(\frac{\Sigma_S}{\sqrt{S}}, \|\cdot\|_{A_D}, \frac{\eta}{\sqrt{S}}\right)\right|\right) d\eta \\ &\leq \sqrt{S}\int_0^\infty \log^{\frac{1}{2}}\left(\left|\mathcal{C}\left(\frac{\Sigma_S}{\sqrt{S}}, \|\cdot\|_{A_D}, \eta\right)\right|\right) d\eta. \end{aligned} \tag{61}$$

An estimate of the covering number involved in (61) is found in the following lemma.

Lemma C.7.3. *For any $\alpha \in \mathbb{C}^{\widetilde{N}}$, consider a norm $\|\cdot\|_A$ on $\mathbb{C}^{\widetilde{N}}$ that satisfies*

$$\|\alpha\|_A = \|A(\alpha)\|_2 \leq \frac{\kappa}{\sqrt{\widetilde{M}}} \cdot \|\alpha\|_1,$$

for some linear map $A(\cdot) : \mathbb{C}^{\widetilde{N}} \to \mathbb{C}^{N'}$ with rank of at most $\widetilde{M}$, $\kappa > 0$, and $N' \in \mathbb{Z}$. Then, for $0 < \eta < \frac{\kappa}{\sqrt{\widetilde{M}}}$ and $\widetilde{M} \gtrsim 1$, we have

$$\log\left(\left|\mathcal{C}\left(\frac{\Sigma_S}{\sqrt{S}}, \|\cdot\|_A, \eta\right)\right|\right) \lesssim \min\left(S\log\widetilde{N} + S\log\left(1 + \frac{2\kappa}{\eta\sqrt{\widetilde{M}}}\right), \frac{\kappa^2}{\eta^2\widetilde{M}} \cdot \log^2\widetilde{N}\right). \tag{62}$$

When $\eta \geq \frac{\kappa}{\sqrt{\widetilde{M}}}$, we have $\left|\mathcal{C}\left(\frac{\Sigma_S}{\sqrt{S}}, \|\cdot\|_A, \eta\right)\right| = 1$.

Proof. See Appendix C.7.5 below. □

Qualitatively speaking, of the two bounds on the right hand of (62), the first is tighter when η is small while the second is more effective for larger values of η. The A_D-norm satisfies the hypothesis of Lemma C.7.3 with $\kappa = \widetilde{\mu}$ and the map $A_D(\cdot)$. Consequently, for some $0 < \alpha \leq \frac{\widetilde{\mu}}{\sqrt{\widetilde{M}}}$ to be set later, we have

$$\begin{aligned}&\int_0^\infty \log^{\frac{1}{2}}\left(\left|\mathcal{C}\left(\frac{\Sigma_S}{\sqrt{S}}, \|\cdot\|_{A_D}, \eta\right)\right|\right) d\eta \\ &= \int_0^\alpha \log^{\frac{1}{2}}\left(\left|\mathcal{C}\left(\frac{\Sigma_S}{\sqrt{S}}, \|\cdot\|_{A_D}, \eta\right)\right|\right) d\eta + \int_\alpha^{\frac{\widetilde{\mu}}{\sqrt{\widetilde{M}}}} \log^{\frac{1}{2}}\left(\left|\mathcal{C}\left(\frac{\Sigma_S}{\sqrt{S}}, \|\cdot\|_{A_D}, \eta\right)\right|\right) d\eta \\ &\lesssim \int_0^\alpha \left(\sqrt{S\log\widetilde{N}} + \sqrt{S\log\left(1 + \frac{2\widetilde{\mu}}{\eta\sqrt{\widetilde{M}}}\right)}\right) d\eta + \log\widetilde{N}\int_\alpha^{\frac{\widetilde{\mu}}{\sqrt{\widetilde{M}}}} \frac{\widetilde{\mu}}{\eta\sqrt{\widetilde{M}}} d\eta \\ &\lesssim \alpha\sqrt{S\log\widetilde{N}} + \alpha\sqrt{S\log\left(1 + \frac{2\widetilde{\mu}}{\alpha\sqrt{\widetilde{M}}}\right)} + \frac{\widetilde{\mu}}{\sqrt{\widetilde{M}}}\log\widetilde{N}\log\left(\frac{\widetilde{\mu}}{\alpha\sqrt{\widetilde{M}}}\right). \end{aligned} \tag{63}$$

The integrals above are calculated via symbolic mathematical programs. With the choice of $\alpha = \frac{\widetilde{\mu}}{\sqrt{S\widetilde{M}}}$, we obtain

$$\begin{aligned}
&\int_0^\infty \log^{\frac{1}{2}}\left(\left|\mathcal{C}\left(\frac{\Sigma_S}{\sqrt{S}}, \|\cdot\|_{A_D}, \eta\right)\right|\right) d\eta \\
&\lesssim \frac{\widetilde{\mu}}{\sqrt{\widetilde{M}}}\sqrt{\log \widetilde{N}} + \frac{\widetilde{\mu}}{\sqrt{\widetilde{M}}}\sqrt{\log(1+2\sqrt{S})} + \frac{\widetilde{\mu}}{\sqrt{\widetilde{M}}}\log S \log \widetilde{N} \\
&\lesssim \frac{\widetilde{\mu}}{\sqrt{\widetilde{M}}}\log S \log \widetilde{N}, \qquad (64)
\end{aligned}$$

for $S \gtrsim 1$. Now putting the estimates in (64) back into (61), we arrive at

$$\gamma_2(\mathcal{A}_D, \|\cdot\|_2) \lesssim \widetilde{\mu}\sqrt{\frac{S}{\widetilde{M}}}\log S \log \widetilde{N}. \qquad (65)$$

Before completing the analysis of the DBD RIP problem, let us calculate the same quantities for RBD case.

C.7.3 Calculating $d_2(\mathcal{A}_R)$, $d_F(\mathcal{A}_R)$, and $\gamma_2(\mathcal{A}_R, \|\cdot\|_2)$

Again, we begin by defining the following A_R-norm for any $\alpha \in \mathbb{C}^{\widetilde{N}}$:

$$\|\alpha\|_{A_R} := \|A_R(\alpha)\|_2 \qquad (66)$$

The A_R-norm can similarly be bounded by the ℓ_1-norm as indicated by the following lemma:

Lemma C.7.4. *For every $\alpha \in \mathbb{C}^{\widetilde{N}}$, it holds that*

$$\|\alpha\|_{A_R} \leq \frac{\gamma}{\sqrt{\widetilde{M}}} \cdot \|\alpha\|_1. \qquad (67)$$

Proof. Note that

$$\begin{aligned}
\|\alpha\|_{A_R} = \|A_R(\alpha)\|_2 &= \frac{1}{\sqrt{M}}\|X_R(\alpha)\|_2 = \frac{1}{\sqrt{M}}\|\sum_{\widetilde{n}\in[\widetilde{N}]} \alpha(\widetilde{n}) X_R(e_{\widetilde{n}})\|_2 \\
&\leq \frac{1}{\sqrt{M}} \sum_{\widetilde{n}\in[\widetilde{N}]} |\alpha(\widetilde{n})| \cdot \|X_R(e_{\widetilde{n}})\|_2 \leq \frac{1}{\sqrt{M}} \max_{\widetilde{n}\in[\widetilde{N}]} \|X_R(e_{\widetilde{n}})\|_2 \cdot \|\alpha\|_1 \\
&= \frac{\gamma}{\sqrt{\widetilde{M}}} \cdot \|\alpha\|_1,
\end{aligned}$$

where we used the linearity of $X_R(\cdot)$, triangle inequality, Hölder inequality, and the definition of γ. $\square$

With the A_R-norm characterized, let us begin computing the quantities involved in Theorem C.7.1. First, we have

$$d_F(\mathcal{A}_R) = \sup_{A_R(\alpha)\in\mathcal{A}_R} \|A_R(\alpha)\|_F = \sup_{\alpha\in\Sigma_S} \|X_R(\alpha)\|_F = \sup_{\alpha\in\Sigma_S} \|x(\alpha)\|_2 = \sup_{\alpha\in\Sigma_S} \|\alpha\|_2 = 1. \tag{68}$$

Second, we have

$$d_2(\mathcal{A}_R) = \sup_{A_R(\alpha)\in\mathcal{A}_R} \|A_R(\alpha)\|_2 = \sup_{\alpha\in\Sigma_S} \|\alpha\|_{A_R} \leq \frac{\gamma}{\sqrt{\widetilde{M}}} \sup_{\alpha\in\Sigma_S} \|\alpha\|_1 \leq \gamma\sqrt{\frac{S}{\widetilde{M}}}, \tag{69}$$

where we used Lemma C.7.4. It remains to bound $\gamma_2(\mathcal{A}_R, \|\cdot\|_2)$. Following the steps for the DBD case, we can write

$$\gamma_2(\mathcal{A}_R, \|\cdot\|_2) \leq \sqrt{S}\int_0^\infty \log^{\frac{1}{2}}\left(\left|\mathcal{C}\left(\frac{\Sigma_S}{\sqrt{S}}, \|\cdot\|_{A_R}, \eta\right)\right|\right) d\eta.$$

Because the A_R-norm satisfies the hypothesis of Lemma C.7.3 with $\kappa = \gamma$ and the map $A_R(\cdot)$, we can follow the same steps for the DBD case to arrive at

$$\gamma_2(\mathcal{A}_R, \|\cdot\|_2) \lesssim \gamma\sqrt{\frac{S}{\widetilde{M}}}\log S \log \widetilde{N}. \tag{70}$$

C.7.4 Denouement

We notice that the quantities $d_F(\mathcal{A}_D), d_2(\mathcal{A}_D)$, and $\gamma_2(\mathcal{A}_D, \|\cdot\|_2)$ have the same bounds as their counterparts $d_F(\mathcal{A}_R), d_2(\mathcal{A}_R)$, and $\gamma_2(\mathcal{A}_R, \|\cdot\|_2)$ except for the type of the coherence factor involved. Therefore, it suffices to complete the proof for the DBD case; the same result will hold for the RBD case modulus a change in the coherence factor.

To begin, given a predetermined conditioning $\delta < 1$, assume that

$$\widetilde{M} \gtrsim_\tau \delta^{-2}\widetilde{\mu}^2 \cdot S\log^2 S \log^2 \widetilde{N}. \tag{71}$$

Equipped with the estimates in Section C.7.2 (i.e. (59), (60), and (65)) we first compute E_1 in Theorem C.7.1:

$$\begin{aligned} E_1 &:= \gamma_2\left(\mathcal{A}_D, \|\cdot\|_2\right)\left(\gamma_2\left(\mathcal{A}_D, \|\cdot\|_2\right) + d_F(\mathcal{A}_D)\right) + d_F(\mathcal{A}_D) d_2(\mathcal{A}_D) \\ &\lesssim \widetilde{\mu}\sqrt{\frac{S}{\widetilde{M}}}\log S \log \widetilde{N}\left(\widetilde{\mu}\sqrt{\frac{S}{\widetilde{M}}}\log S\log\widetilde{N}+1\right) + \widetilde{\mu}\sqrt{\frac{S}{\widetilde{M}}} \\ &\lesssim_\tau \delta\left(\delta+1\right) + \frac{\delta}{\log S \log \widetilde{N}} \\ &\lesssim \delta, \end{aligned}$$

where we assumed that $S \gtrsim 1$ and used the fact that $\delta < 1$ in the last line. Moving on to E_2, we obtain

$$\begin{aligned} E_2 &:= d_2(\mathcal{A}_D)\left(\gamma_2\left(\mathcal{A}_D, \|\cdot\|_2\right) + d_F(\mathcal{A}_D)\right) \\ &\lesssim \widetilde{\mu}\sqrt{\frac{S}{\widetilde{M}}}\left(\widetilde{\mu}\sqrt{\frac{S}{\widetilde{M}}}\log S\log\widetilde{N}+1\right) \\ &\lesssim_\tau \frac{\delta}{\log S\log\widetilde{N}}\left(\delta+1\right) \\ &\lesssim \frac{\delta}{\log S\log\widetilde{N}}. \end{aligned}$$

Similarly for E_3, we obtain

$$E_3 := d_2^2(\mathcal{A}_D) \leq \frac{\widetilde{\mu}^2 S}{\widetilde{M}} \lesssim_\tau \frac{\delta^2}{\log^2 S \log^2 \widetilde{N}}.$$

Applying the above estimates of E_1, E_2, and E_3 into the tail bound in Theorem C.7.1, we obtain

$$\log \mathbb{P}\left\{\sup_{\alpha\in\Sigma_S}\left|\|\Phi\cdot x(\alpha)\|_2^2 - 1\right| \gtrsim_\tau \delta + t\right\} \lesssim_\tau -\min\left(\delta^{-2}t^2\log^2 S\log^2\widetilde{N}, \delta^{-1}t\log^2 S\log^2\widetilde{N}\right).$$

Substituting $\epsilon = \delta$ and assuming that $S \gtrsim 1$, we arrive at

$$\begin{aligned} &\log \mathbb{P}\left\{\sup_{\alpha\in\Sigma_S}\left|\|\Phi\cdot x(\alpha)\|_2^2 - 1\right| \gtrsim_\tau \delta\right\} \\ &\qquad\qquad \lesssim_\tau -\log^2 S\log^2\widetilde{N}. \end{aligned}$$

This in turn implies that

$$\log \mathbb{P}\left\{ \sup_{\alpha \in \Sigma_S} \left| \|\Phi \cdot x(\alpha)\|_2^2 - 1 \right| > \delta \right\} \lesssim_\tau - \log^2 S \log^2 \widetilde{N},$$

which completes the proof of Theorem 3.2.1. Replacing $\widetilde{\mu}$ with γ and repeating this argument concludes the proof of Theorem 3.2.2 as well.

C.7.5 Proof of Lemma C.7.3

To bound the covering numbers of $\frac{U(\Sigma_S)}{\sqrt{S}}$, we use arguments adapted from [118]. First, we have

$$\frac{U\left(\Sigma_S\right)}{\sqrt{S}} = \bigcup_{|T|=S} \frac{\mathcal{B}_2^T}{\sqrt{S}}, \tag{72}$$

where $\mathcal{B}_2^T$ denote the ℓ_2-unit ball in $\mathbb{C}^T$ (i.e., the subspace spanned by $\{e_{\widetilde{n}}\}_{\widetilde{n} \in T}$). Next, notice that when $\alpha \in \frac{U(\Sigma_S)}{\sqrt{S}}$, then $\|\alpha\|_1 \leq 1$. Subsequently, from our assumptions about the A-norm in Lemma C.7.3, we have

$$\|\alpha\|_A \leq \frac{\kappa}{\sqrt{\widetilde{M}}}. \tag{73}$$

This implies that for every support $T \subset [\widetilde{N}]$ with $|T| = S$, we have

$$\frac{\mathcal{B}_2^T}{\sqrt{S}} \subseteq \frac{\kappa}{\sqrt{\widetilde{M}}} \mathcal{B}_A^T, \tag{74}$$

where $\mathcal{B}_A^T$ denote the unit ball with respect to the A-norm in $\mathbb{C}^T$. Together, (72) and (74) imply that

$$\frac{U\left(\Sigma_S\right)}{\sqrt{S}} \subseteq \bigcup_{|T|=S} \frac{\kappa}{\sqrt{\widetilde{M}}} \mathcal{B}_A^T. \tag{75}$$

On the other hand, we have

$$\frac{U\left(\Sigma_S\right)}{\sqrt{S}} \subseteq \frac{\kappa}{\sqrt{\widetilde{M}}} \mathcal{B}_A^{\widetilde{N}},$$

where $\mathcal{B}_A^{\widetilde{N}}$ is the unit ball measured with the A-norm in $\mathbb{C}^{\widetilde{N}}$, which implies that

$$\left| \mathcal{C}\left(\frac{U\left(\Sigma_S\right)}{\sqrt{S}}, \|\cdot\|_A, u \right) \right| = 1,$$

if $u \geq \frac{\kappa}{\sqrt{\widetilde{M}}}$. This proves the second statement in Lemma C.7.3.

Otherwise if $u < \frac{\kappa}{\sqrt{\widetilde{M}}}$, we use the containment given in (75) to deduce that

$$\begin{aligned}\left|\mathcal{C}\left(\frac{U(\Sigma_S)}{\sqrt{S}}, \|\cdot\|_A, u\right)\right| &\leq \left|\mathcal{C}\left(\bigcup_{|T|=S} \frac{\kappa}{\sqrt{\widetilde{M}}}\mathcal{B}_A^T, \|\cdot\|_A, u\right)\right| \\ &\leq \binom{\widetilde{N}}{S}\left|\mathcal{C}\left(\frac{\kappa}{\sqrt{\widetilde{M}}}\mathcal{B}_A^T, \|\cdot\|_A, u\right)\right| \\ &= \binom{\widetilde{N}}{S}\left|\mathcal{C}\left(\mathcal{B}_A^T, \|\cdot\|_A, \frac{\sqrt{\widetilde{M}}}{\kappa}u\right)\right|. \end{aligned} \tag{76}$$

A bound on the covering number on the right-hand side of (76) follows closely that given in Lemma 2.5.2 (see also [136, Lemma 5.2]), and it is given by

$$\mathcal{C}\left(\mathcal{B}_A^T, \|\cdot\|_A, u\right) \leq \left(1 + \frac{2}{u}\right)^{2|T|}.$$

Using this bound on (76), we have

$$\left|\mathcal{C}\left(\frac{U(\Sigma_S)}{\sqrt{S}}, \|\cdot\|_A, u\right)\right| \leq \binom{\widetilde{N}}{S}\left(1 + \frac{2\kappa}{u\sqrt{\widetilde{M}}}\right)^{2S} \leq \left(\frac{e\widetilde{N}}{S}\right)^{S}\left(1 + \frac{2\kappa}{u\sqrt{\widetilde{M}}}\right)^{2S},$$

where the last inequality follows from Stirling's inequality. Thus,

$$\begin{aligned}\log\left(\left|\mathcal{C}\left(\frac{U(\Sigma_S)}{\sqrt{S}}, \|\cdot\|_A, u\right)\right|\right) &\leq S\log\left(e\widetilde{N}\right) + 2S\log\left(1 + \frac{2\kappa}{u\sqrt{\widetilde{M}}}\right) \\ &\lesssim S\log\widetilde{N} + S\log\left(1 + \frac{2\kappa}{u\sqrt{\widetilde{M}}}\right). \end{aligned} \tag{77}$$

The above bound is the first of the two bounds that we want to prove. This bound works fine when the resolution u is small, but when u is large, we require a better covering number for $\frac{U(\Sigma_S)}{\sqrt{S}}$ to achieve a sharper bound.

To achieve this bound, we start from the following containment

$$\frac{U(\Sigma_S)}{\sqrt{S}} \subseteq \mathcal{B}_1^{\widetilde{N}},$$

where $\mathcal{B}_1^{\widetilde{N}}$ is the unit ℓ_1-ball in $\mathbb{C}^{\widetilde{N}}$. This containment allows us to write

$$\left|\mathcal{C}\left(\frac{U(\Sigma_S)}{\sqrt{S}}, \|\cdot\|_A, u\right)\right| \leq \left|\mathcal{C}\left(\mathcal{B}_1^{\widetilde{N}}, \|\cdot\|_A, u\right)\right|. \tag{78}$$

To bound the covering number above, first let $B_1(\mathbb{R}^{\widetilde{N}})$ denote the ℓ_1-ball in $\mathbb{R}^{\widetilde{N}}$ (recall $\mathcal{B}_1^{\widetilde{N}}$ the equivalent ℓ_1-ball but in $\mathbb{C}^{\widetilde{N}}$). Next, consider an arbitrary $\beta \in B_1(\mathbb{R}^{\widetilde{N}})$ and consider a random vector Z that takes value $\operatorname{sgn}(\beta_{\widetilde{n}})e_{\widetilde{n}}$ with probability $|\beta_{\widetilde{n}}|$ for $\widetilde{n} \in [\widetilde{N}]$.[2] Clearly, $\mathbb{E}\{Z\} = \beta$. Now, we wish to approximate β with the average of L independent copies of Z denoted by $\{Z_l\}$ for $l \in [L]$. The expected approximation error measured in the A-norm is

$$\mathbb{E}\left\{\left\|\beta - \frac{1}{L}\sum_{l\in[L]} Z_l\right\|_A\right\}.$$

Since the argument of the norm is centered at zero, we can use the symmetrization technique (Lemma A.1.1) to yield

$$\mathbb{E}\left\{\left\|\beta - \frac{1}{L}\sum_{l\in[L]} Z_l\right\|_A\right\} = \frac{1}{L}\mathbb{E}\left\{\left\|\sum_{l\in[L]} Z_l - \mathbb{E}\{Z_l\}\right\|_A\right\} \leq \frac{2}{L}\mathbb{E}\left\{\left\|\sum_{l\in[L]} \xi_l Z_l\right\|_A\right\}. \tag{79}$$

Now, by linearity of the function $A(\cdot)$, we have

$$\mathbb{E}\left\{\left\|\sum_{l\in[L]} \xi_l Z_l\right\|_A\right\} = \mathbb{E}\left\{\left\|\sum_{l\in[L]} \xi_l A(Z_l)\right\|_2\right\},$$

where by assumption in Lemma C.7.3, $A(Z_l) \in \mathbb{C}^{\widetilde{M}\times N'}$ is of rank at most $\widetilde{M}$ for every l. Using Khintchine's inequality for matrices (Lemma A.1.6) and Jensen's inequality,

[2] We use the convention that $\operatorname{sgn}(0) = 0$.

we continue the argument as follows:

$$\begin{aligned}
\mathbb{E}\left\{\left\|\sum_{l\in[L]}\xi_l A(Z_l)\right\|_2\right\} &= \mathbb{E}_Z\left\{\mathbb{E}_\xi\left\{\left\|\sum_{l\in[L]}\xi_l A(Z_l)\right\|_2\right\}\right\} \\
&\lesssim \sqrt{\log\widetilde{M}}\cdot\mathbb{E}_Z\left\{\left(\sum_{l\in[L]}\|A(Z_l)\|_2^2\right)^{1/2}\right\} \\
&\lesssim \sqrt{\log\widetilde{M}}\cdot\left(\sum_{l\in[L]}\mathbb{E}\left\{\|A(Z_l)\|_2^2\right\}\right)^{1/2} \\
&\lesssim \sqrt{\log\widetilde{M}}\cdot\sqrt{\sum_{l\in[L]}\sum_{\widetilde{n}\in[\widetilde{N}]}|\beta_{\widetilde{n}}|\cdot\|A(e_{\widetilde{n}})\|_2^2} \\
&\lesssim \sqrt{\log\widetilde{M}}\cdot\sqrt{L\max_{\widetilde{n}\in[\widetilde{N}]}\|A(e_{\widetilde{n}})\|_2^2} \\
&= \frac{\kappa\sqrt{L\log\widetilde{M}}}{\sqrt{\widetilde{M}}}. \qquad (80)
\end{aligned}$$

Using the above inequality in combination with (79) yields

$$\mathbb{E}\left\{\left\|\beta-\frac{1}{L}\sum_{l\in[L]}Z_l\right\|_A\right\}\lesssim\frac{\kappa\sqrt{\log\widetilde{M}}}{\sqrt{L\widetilde{M}}}.$$

To keep the average no larger than u, it suffices to take

$$L\gtrsim\frac{\kappa^2\log\widetilde{M}}{u^2\widetilde{M}}.$$

With this choice of L, there exists a linear combination of independent copies of Z that falls within a u distance of β. In other words, we have shown that for an arbitrary $\beta\in B_1(\mathbb{R}^{\widetilde{N}})$, there exists an average of L elements of $\{\pm e_{\widetilde{N}}\}\cup\{0\}$ that is of distance at most u from β. There are $2\widetilde{N}+1$ elements in the aforementioned set, and so $(2\widetilde{N}+1)^L$ possibilities for the average. We therefore conclude that

$$\begin{aligned}
\log\left(\left|\mathcal{C}\left(B_1(\mathbb{R}^{\widetilde{N}}),\|\cdot\|_A,u\right)\right|\right) &\lesssim \log\widetilde{M}\cdot\frac{\kappa^2}{u^2\widetilde{M}}\cdot\log\left(2\widetilde{N}+1\right) \\
&\lesssim \log\widetilde{M}\cdot\frac{\kappa^2}{u^2\widetilde{M}}\log\widetilde{N} \\
&\leq \frac{\kappa^2}{u^2\widetilde{M}}\log^2\widetilde{N}, \qquad (81)
\end{aligned}$$

when $\widetilde{N} \gtrsim 1$. Now consider an arbitrary $\beta \in \mathcal{B}_1^{\widetilde{N}}$ (complex ball), and note that $\operatorname{Re}\{\beta\}, \operatorname{Im}\{\beta\} \in B_1(\mathbb{R}^{\widetilde{N}})$ (real ball). Let $\mathcal{C}_1 := \mathcal{C}\left(B_1(\mathbb{R}^{\widetilde{N}}), \|\cdot\|_A, \frac{u}{2}\right)$ denote the minimal $(\frac{u}{2}, \|\cdot\|_A)$-cover for $B_1(\mathbb{R}^{\widetilde{N}})$. Therefore, there exist $p_1, p_2 \in \mathcal{C}_1$ such that $\|\operatorname{Re}\{\beta\} - p_1\|_A \leq \frac{u}{2}$ and $\|\operatorname{Im}\{\beta\} - p_2\|_A \leq \frac{u}{2}$. It follows that

$$\begin{aligned}\|\beta - (p_1 + jp_2)\|_A &= \|(\operatorname{Re}\{\beta\} - p_1) + j(\operatorname{Im}\{\beta\} - p_2)\|_A \\ &\leq \|\operatorname{Re}\{\beta\} - p_1\|_A + \|\operatorname{Im}\{\beta\} - p_2\|_A \\ &\leq u.\end{aligned}$$

Therefore, $\{p_1 + jp_2 : p_1,\, p_2 \in \mathcal{C}_1\}$ is a cover for $\mathcal{B}_1^{\widetilde{N}}$, and clearly,

$$\left|\mathcal{C}\left(\mathcal{B}_1^{\widetilde{N}}, \|\cdot\|_A, u\right)\right| \leq \left|\mathcal{C}\left(B_1(\mathbb{R}^{\widetilde{N}}), \|\cdot\|_A, \frac{u}{2}\right)\right|^2.$$

It now follows from (78) and (81) that

$$\log\left(\left|\mathcal{C}\left(\frac{U(\Sigma_S)}{\sqrt{S}}, \|\cdot\|_A, u\right)\right|\right) \lesssim \frac{\kappa^2}{u^2\widetilde{M}} \log^2 \widetilde{N} \tag{82}$$

Combining (77) and (82), we finally arrive at

$$\log\left(\left|\mathcal{C}\left(\frac{U(\Sigma_S)}{\sqrt{S}}, \|\cdot\|_A, u\right)\right|\right) \lesssim \min\left\{S\log\widetilde{N} + S\log\left(1 + \frac{2\kappa}{u\sqrt{\widetilde{M}}}\right), \log^2\widetilde{N} \cdot \frac{\kappa^2}{u^2\widetilde{M}}\right\}.$$

APPENDIX D

STABLE MANIFOLD EMBEDDING PROOF

In this appendix, we prove Theorem 4.1.2 of Chapter 4. Suppose $\mathcal{M}$ is the Riemannian submanifold considered in Theorem 4.1.2 and define the set of *chords* of $\mathcal{M}$ (i.e., the set of all normalized difference vectors in $\mathcal{M}$) as

$$U(\mathcal{M}-\mathcal{M}) = \left\{ \frac{x-y}{\|x-y\|_2} \;\middle|\; x,y \in \mathcal{M},\ x \neq y \right\}.$$

Then, $\widehat{\Phi}$ provides a stable embedding of $\mathcal{M}$ with conditioning $\delta_{\mathcal{M}}$ if and only if $\sup_{x \in U(\mathcal{M}-\mathcal{M})} \left| \|\widehat{\Phi}x\|_2^2 - 1 \right| \leq \delta_{\mathcal{M}}$. In other words, $\widehat{\Phi}$ provides a stable embedding of $\mathcal{M}$ if and only if $\widehat{\Phi}$ approximately preserves the norms of all elements in $U(\mathcal{M}-\mathcal{M})$. We will use this equivalence for the proof of Theorem 4.1.2.

The proof of Theorem 4.1.2 follows very closely the proof technique of [15] and is basically comprised of three steps. The first step involves judiciously choosing a *generalized* covering set B of the manifold $\mathcal{M}$ using a collection of points on the manifold and their corresponding tangent planes. Lemma D.1.1 then shows that every point of $U(\mathcal{M}-\mathcal{M})$ can be approximated by some point in $U(B-B)$. The second step (encapsulated by Lemma D.2.1) then applies the JL lemma for RIP operators (i.e., Theorem 4.1.1) to obtain an approximate norm preservation of all elements of $U(B-B)$. Finally in Section D.3, we extend this approximate norm preservation to all points on $U(\mathcal{M}-\mathcal{M})$ via simple geometric arguments. As described, the proof technique here distinguishes from that of [15] mainly in the separation of the stable embedding operator from the covering of the manifold.

D.1 Covering $U(\mathcal{M}-\mathcal{M})$

In this section, we construct a set B and show in Lemma D.1.1 that $U(B-B)$ is a cover of $U(\mathcal{M}-\mathcal{M})$. Let $A = A(T) := \mathcal{C}(\mathcal{M}, d_{\mathcal{M}}, T)$ for some $T \leq \tau$ be the $(T, d_{\mathcal{M}})$-cover of $\mathcal{M}$ of minimum cardinality. For any $x \in \mathcal{M}$, we can find an $a \in A$ such that $d_{\mathcal{M}}(a, x) \leq T$. Define a generalized covering set B of the manifold $\mathcal{M}$ as

$$B = B(T) = \bigcup_{a \in A} \{a + \mathcal{T}_a\mathcal{M}(T)\},$$

where $\mathcal{T}_a\mathcal{M}(T) := \{u \in \mathcal{T}_a\mathcal{M} \mid \|u\|_2 \leq T\}$ refers to all tangent vectors of $\mathcal{M}$ at the point a whose lengths are less than T. B is called a generalized covering set as it is a union of (subsets of) affine D-dimensional planes of $\mathbb{R}^N$ (i.e., B is not a finite set).

The goal of this section is to show that $U(B-B)$ is a suitable cover of $U(\mathcal{M}-\mathcal{M})$ as detailed in the following lemma:

Lemma D.1.1. *Let B be defined as above. For $T \leq \tau$, set $\epsilon(T) := 4\sqrt{\frac{T}{\tau}}$. Then, $U(B-B)$ is an $(\epsilon(T), \|\cdot\|_2)$-cover of $U(\mathcal{M}-\mathcal{M})$. In other words, for every $u \in U(\mathcal{M}-\mathcal{M})$, we can find a $b \in U(B-B)$ such that $\|u-b\|_2 \leq \epsilon(T)$.*

Proof. To prove this lemma, we break the set of chords $U(\mathcal{M}-\mathcal{M})$ into sets of "long" and "short" chords which we will cover separately. The sets of short and long chords (delineated by Euclidean distance $\frac{T}{2}$) are defined by:

$$\begin{aligned} U^s(\mathcal{M}-\mathcal{M}) &:= \left\{ \frac{x_1-x_2}{\|x_1-x_2\|_2} \mid x_1, x_2 \in \mathcal{M}, 0 < \|x_1-x_2\|_2 \leq \frac{T}{2} \right\}, \\ U^l(\mathcal{M}-\mathcal{M}) &:= \left\{ \frac{x_1-x_2}{\|x_1-x_2\|_2} \mid x_1, x_2 \in \mathcal{M}, \|x_1-x_2\|_2 > \frac{T}{2} \right\}, \end{aligned}$$

and $U(\mathcal{M}-\mathcal{M}) = U^s(\mathcal{M}-\mathcal{M}) \cup U^l(\mathcal{M}-\mathcal{M})$.

Let us start with the cover of $U^s(\mathcal{M}-\mathcal{M})$ where, due to the locally Euclidean structure of manifolds, the short chords in $U^s(\mathcal{M}-\mathcal{M})$ can be approximated by tangent vectors of the manifold. Pick an element $\frac{x_1-x_2}{\|x_1-x_2\|_2}$ of $U^s(\mathcal{M}-\mathcal{M})$ where by definition $\|x_1-x_2\|_2 \leq \frac{T}{2}$. From Lemma 2.5.1, $\|x_1-x_2\|_2 \leq \frac{T}{2} \leq \frac{\tau}{2}$ (since we assume

$T \leq \tau$) implies that

$$d_{\mathcal{M}}(x_1, x_2) \leq \tau - \tau\sqrt{1 - \frac{2\|x_1 - x_2\|_2}{\tau}} \leq \tau - \tau\left(1 - \frac{2\|x_1 - x_2\|_2}{\tau}\right) = 2\|x_1 - x_2\|_2 \leq T. \quad (83)$$

Now, let $\mu := d_{\mathcal{M}}(x_1, x_2)$ and let $\gamma(t)$ be the unit-speed geodesic parameterization from x_1 to x_2 where $\gamma(0) = x_1$, $\gamma(\mu) = x_2$, and $\gamma'(0) \in U(\mathcal{T}_{x_1}\mathcal{M})$. From Lemma 2.5.1, we have

$$x_1 - x_2 = \gamma(\mu) - \gamma(0) = \mu\gamma'(0) + r, \quad (84)$$

with $\|r\|_2 \leq \frac{\mu^2}{2\tau}$. Let $a \in A$ be the closest geodesic covering point to x_1 (so that $d_{\mathcal{M}}(a, x_1) \leq T$) and let $b \in U(\mathcal{T}_a\mathcal{M})$ be the parallel transport of $\gamma'(0)$ onto $\mathcal{T}_a\mathcal{M}$. First, $b \in U(B - B)$ by definition of the set B and second, Lemma 2.5.1 says that $\langle\gamma'(0), b\rangle \geq 1 - \frac{d_{\mathcal{M}}(a,x_1)}{\tau} \geq 1 - \frac{T}{\tau}$, since $d_{\mathcal{M}}(a, x_1) \leq T \leq \tau$. Thus,

$$\|\gamma'(0) - b\|_2^2 = \|\gamma'(0)\|_2^2 + \|b\|_2^2 - 2\langle\gamma'(0), b\rangle = 2(1 - \langle\gamma'(0), b\rangle) \leq \frac{2T}{\tau}. \quad (85)$$

We now show that b is indeed close to the short chord $\frac{x_1 - x_2}{\|x_1 - x_2\|_2}$ by combining (84) and (85):

$$\begin{aligned}\left\|\frac{x_1 - x_2}{\|x_1 - x_2\|_2} - b\right\|_2 &= \left\|\frac{\mu\gamma'(0) + r}{\|x_1 - x_2\|_2} - b\right\|_2 \\ &= \left\|\gamma'(0) - b + \left(\frac{\mu}{\|x_1 - x_2\|_2} - 1\right)\gamma'(0) + \frac{r}{\|x_1 - x_2\|_2}\right\|_2 \\ &\leq \sqrt{\frac{2T}{\tau}} + \left(\frac{\mu}{\|x_1 - x_2\|_2} - 1\right) + \frac{\mu}{\|x_1 - x_2\|_2} \cdot \frac{\mu}{2\tau}. \quad (86)\end{aligned}$$

To remove the dependence of (86) on $\|x_1 - x_2\|_2$, we use Lemma 2.5.1 to obtain

$$\|x_1 - x_2\|_2 \geq \mu - \frac{\mu^2}{2\tau} = \mu\left(1 - \frac{\mu}{2\tau}\right) \Leftrightarrow \frac{\mu}{\|x_1 - x_2\|_2} \leq \frac{1}{\left(1 - \frac{\mu}{2\tau}\right)} \leq 1 + \frac{\mu}{\tau},$$

where we used the inequality $\frac{1}{1-a} \leq (1 + 2a)$ whenever $0 < a \leq \frac{1}{2}$ (this is true since $\frac{\mu}{2\tau} \leq \frac{T}{2\tau} \leq \frac{1}{2}$). Applying this to (86) and applying $\mu \leq T$ obtained earlier in (83), we obtain

$$\left\|\frac{x_1 - x_2}{\|x_1 - x_2\|_2} - b\right\|_2 \leq \sqrt{\frac{2T}{\tau}} + \frac{\mu}{\tau} + \left(1 + \frac{\mu}{\tau}\right)\frac{\mu}{2\tau} \leq 4\sqrt{\frac{T}{\tau}} =: \epsilon_1(T),$$

where we used the fact that $\frac{T^2}{\tau^2} \leq \frac{T}{\tau} \leq \sqrt{\frac{T}{\tau}} \leq 1$. This proves that for every element of $U^s(\mathcal{M}-\mathcal{M})$, we can find an element $b \in U(B-B)$ that is within $\epsilon_1(T)$ of it. Thus, $U(B-B)$ is an $(\epsilon_1(T), \|\cdot\|_2)$-cover of $U^s(\mathcal{M}-\mathcal{M})$.

Let us now move on to covering $U^l(\mathcal{M}-\mathcal{M})$. Pick an element $\frac{x_1-x_2}{\|x_1-x_2\|_2}$ of $U^l(\mathcal{M}-\mathcal{M})$. For each x_i, for $i=1,2$, choose its closest geodesic covering point $a_i \in A$ so that $\mu_i := d_{\mathcal{M}}(a_i, x_i) \leq T$. Let $\gamma_i(t)$ be the unit-speed geodesic parameterization from a_i to x_i, so that $\gamma_i(0) = a_i$, $\gamma_i(\mu_i) = x_i$, and $\gamma_i'(0) \in U(\mathcal{T}_{a_i}\mathcal{M})$. From Lemma 2.5.1, we have $x_i - a_i = \gamma_i(\mu_i) - \gamma_i(0) = \mu_i\gamma_i'(0) + r_i$, with $\|r_i\|_2 \leq \frac{\mu_i^2}{2\tau}$. Define $b_i = a_i + \mu_i\gamma_i'(0)$ where it is clear that $x_i - b_i = r_i$ and $b_i \in \{a_i + \mathcal{T}_{a_i}\mathcal{M}(T)\} \subset B$. We will use $\frac{b_1-b_2}{\|b_1-b_2\|_2} \in U(B-B)$ as a covering point near to $\frac{x_1-x_2}{\|x_1-x_2\|_2}$. Following [39], we have

$$\begin{aligned}
& \left\| \frac{x_1-x_2}{\|x_1-x_2\|_2} - \frac{b_1-b_2}{\|b_1-b_2\|_2} \right\|_2 \\
= & \left\| \frac{(x_1-x_2)-(b_1-b_2)}{\|x_1-x_2\|_2} + \frac{(b_1-b_2)(\|b_1-b_2\|_2 - \|x_1-x_2\|_2)}{\|x_1-x_2\|_2\|b_1-b_2\|_2} \right\|_2 \\
\leq & \left\| \frac{(x_1-x_2)-(b_1-b_2)}{\|x_1-x_2\|_2} \right\|_2 + \left\| \frac{(b_1-b_2)(\|b_1-b_2\|_2 - \|x_1-x_2\|_2)}{\|x_1-x_2\|_2\|b_1-b_2\|_2} \right\|_2 .
\end{aligned}$$

We will calculate each of the terms separately. For the first term, we see that

$$\begin{aligned}
\left\| \frac{(x_1-x_2)-(b_1-b_2)}{\|x_1-x_2\|_2} \right\|_2 &= \left\| \frac{(x_1-b_1)-(x_2-b_2)}{\|x_1-x_2\|_2} \right\|_2 \\
&\leq \frac{1}{\|x_1-x_2\|_2}\left(\frac{\mu_1^2}{2\tau} + \frac{\mu_2^2}{2\tau}\right) \\
&\leq \frac{T^2}{\tau\|x_1-x_2\|_2}.
\end{aligned}$$

For the second term, we have

$$\begin{aligned}
\left\| \frac{(b_1-b_2)(\|b_1-b_2\|_2 - \|x_1-x_2\|_2)}{\|x_1-x_2\|_2\|b_1-b_2\|_2} \right\|_2 &= \frac{|\|x_1-x_2\|_2 - \|b_1-b_2\|_2|}{\|x_1-x_2\|_2} \\
&\leq \frac{\|(x_1-x_2)-(b_1-b_2)\|_2}{\|x_1-x_2\|_2} \\
&= \frac{\|(x_1-b_1)-(x_2-b_2)\|_2}{\|x_1-x_2\|_2} \\
&\leq \frac{T^2}{\tau\|x_1-x_2\|_2},
\end{aligned}$$

where we used the reverse triangle inequality in the second line. Now, our definition of long chords implies that $\|x_1 - x_2\|_2 > \frac{T}{2}$. Therefore,

$$\left\| \frac{x_1 - x_2}{\|x_1 - x_2\|_2} - \frac{b_1 - b_2}{\|b_1 - b_2\|_2} \right\|_2 \quad < \quad \frac{4T}{\tau} =: \epsilon_2(T).$$

Thus, $U(B - B)$ is an $(\epsilon_2(T), \|\cdot\|_2)$-cover of $U^l(\mathcal{M} - \mathcal{M})$.

Putting everything together, since $\epsilon_1(T) = 4\sqrt{\frac{T}{\tau}} \geq 4\frac{T}{\tau} = \epsilon_2(T)$, we have that $U(B - B)$ is a $(4\sqrt{\frac{T}{\tau}}, \|\cdot\|_2)$-cover of $U(\mathcal{M} - \mathcal{M})$. □

D.2 Applying the JL Lemma

We want to use $U(B - B)$ as a proxy for $U(\mathcal{M} - \mathcal{M})$ for applying Theorem 4.1.1. However, $U(B - B)$ is not just a finite collection of points and thus Theorem 4.1.1 cannot be applied directly. Fortunately, it is well-known that unit spheres on planes (or affine planes) can be well-covered by a finite collection of points. Indeed, as a corollary to Lemma 2.5.2 (see also [111]), the $(\epsilon, \|\cdot\|_2)$-covering number of a D-dimensional sphere is $\left(1 + \frac{2}{\epsilon}\right)^D$.

$U(B - B)$ can be divided into two sets of elements, namely:

1. $B_1 := U\left(\bigcup_{a \in A} \{\mathcal{T}_a \mathcal{M}(T) - \mathcal{T}_a \mathcal{M}(T)\}\right) = U\left(\bigcup_{a \in A} \mathcal{T}_a \mathcal{M}\right)$, and
2. $B_2 := U\left(\bigcup_{a_1, a_2 \in A, a_1 \neq a_2} \{(a_1 - a_2) + (\mathcal{T}_{a_1} \mathcal{M}(T) - \mathcal{T}_{a_2} \mathcal{M}(T))\}\right)$.

The set B_1 is comprised of $|A|$ D-dimensional unit spheres. From our earlier discussion, we know that each unit sphere can be $(\epsilon, \|\cdot\|_2)$-covered by at most $(1 + \frac{2}{\epsilon})^D$ points. Thus, $|\mathcal{C}(B_1, \|\cdot\|_2, \epsilon)| \leq |A|(1 + \frac{2}{\epsilon})^D$. The set B_2 is the projection onto the unit sphere (in $\mathbb{R}^N$) of not more than $|A|^2$ subsets of affine planes where each affine plane is contained in a linear subspace of dimension $2D + 1$. Thus, $|\mathcal{C}(B_2, \|\cdot\|_2, \epsilon)| \leq |A|^2(1 + \frac{2}{\epsilon})^{2D+1}$.

Define the collection of points $E(\epsilon) := \mathcal{C}(B_1, \|\cdot\|_2, \epsilon) \cup \mathcal{C}(B_2, \|\cdot\|_2, \epsilon)$. From our previous discussion, the cardinality of $E(\epsilon)$ is bounded by

$$|E(\epsilon)| \leq |A| \left(1 + \frac{2}{\epsilon}\right)^D + |A|^2 \left(1 + \frac{2}{\epsilon}\right)^{2D+1} \leq 2|A|^2 \left(1 + \frac{2}{\epsilon}\right)^{2D+1}. \tag{87}$$

By construction, for any $b \in U(B-B)$, we can find an $e \in E(\epsilon)$ such that $\|b-e\|_2 \leq \epsilon$. With the aid of $E(\epsilon)$, we can show the stable embedding of $U(B-B)$ by the operator $\widehat{\Phi}$ defined in Theorem 4.1.2.

Lemma D.2.1. *Choose any failure probability ρ and conditioning $\delta'_{\mathcal{M}} \leq \frac{4}{9}$. Set the covering resolution ϵ in the set $E(\epsilon)$ to $\epsilon = \frac{\delta'_{\mathcal{M}}}{N+1}$. Suppose we have a matrix Φ satisfying the RIP of order $S \geq 40 \log\left(\frac{4|E(\epsilon)|}{\rho}\right)$ and conditioning $\delta \leq \frac{\delta'_{\mathcal{M}}}{4}$. Then, with probability exceeding $1-\rho$, the matrix $\widehat{\Phi} := \Phi D_\xi \Psi$ is a (non-squared) stable embedding[1] of B with conditioning $\frac{9}{4}\delta'_{\mathcal{M}}$ (i.e., $\sup_{b \in U(B-B)} \left| \|\widehat{\Phi} b\|_2 - 1 \right| \leq \frac{9}{4}\delta'_{\mathcal{M}}$).*

Proof of Lemma D.2.1. Fix $\rho < 1$ and $\delta'_{\mathcal{M}} \leq \frac{4}{9}$. If Φ satisfies RIP-(S, δ) with $S \geq 40 \log\left(\frac{4|E(\epsilon)|}{\rho}\right)$ (with ϵ to be defined later) and $\delta \leq \frac{\delta'_{\mathcal{M}}}{4}$ as assumed in Lemma D.2.1, then Theorem 4.1.1 states that with probability exceeding $1-\rho$, $\sup_{e \in E(\epsilon)} \left| \|\widehat{\Phi} e\|_2 - 1 \right| \leq \sup_{e \in E(\epsilon)} \left| \|\widehat{\Phi} e\|_2^2 - 1 \right| \leq \delta'_{\mathcal{M}}$. For a fixed $b \in U(B-B)$, find its nearest covering point $e \in E(\epsilon)$ such that $\|b-e\|_2 \leq \epsilon$. Then, we have

$$\|\widehat{\Phi} b\|_2 \leq \|\widehat{\Phi} e\|_2 + \left\| \widehat{\Phi}(b-e) \right\|_2 \leq (1+\delta'_{\mathcal{M}}) + \|\Phi\|_2 \left\| D_\xi \Psi (b-e) \right\|_2 . \tag{88}$$

Now, it is easy to show that for a matrix $\Phi \in \mathbb{C}^{M \times N}$ that satisfies RIP-(S, δ), $\|\Phi\|_2 \leq \left(\frac{N}{S} + 1\right)(1+\delta)$. Applying this fact to (88), we have

$$\|\widehat{\Phi} b\|_2 \leq (1+\delta'_{\mathcal{M}}) + \left(\frac{N}{S} + 1\right)(1+\delta)\epsilon \leq (1+\delta'_{\mathcal{M}}) + (N+1)\left(1 + \frac{\delta'_{\mathcal{M}}}{4}\right)\epsilon.$$

To remove the catastrophic dependence on $(N+1)$, set $\epsilon = \frac{\delta'_{\mathcal{M}}}{N+1}$. Using this choice of ϵ, we have

$$\|\widehat{\Phi} b\|_2 \leq (1+\delta'_{\mathcal{M}}) + \left(1 + \frac{\delta'_{\mathcal{M}}}{4}\right)\delta'_{\mathcal{M}} \leq 1 + 2\delta'_{\mathcal{M}} + \frac{(\delta'_{\mathcal{M}})^2}{4} \leq 1 + \frac{9}{4}\delta'_{\mathcal{M}},$$

[1] *Squared and non-squared stable embeddings differ only by a small constant in their conditioning. To be more concrete, suppose $C \subset \mathbb{R}^N$. Then it can be shown that $\sup_{c \in C} \left| \|\widehat{\Phi} c\|_2 - 1 \right| \leq \sup_{c \in C} \left| \|\widehat{\Phi} c\|_2^2 - 1 \right|$. Furthermore if $\sup_{c \in C} \left| \|\widehat{\Phi} c\|_2 - 1 \right| \leq 1$, then it can be shown that $\sup_{c \in C} \left| \|\widehat{\Phi} c\|_2^2 - 1 \right| \leq 3 \sup_{c \in C} \left| \|\widehat{\Phi} c\|_2 - 1 \right|$.*

where we used the fact that $\delta'_{\mathcal{M}} \leq \frac{4}{9} \leq 1$. Using the same steps for the lower bound, we obtain

$$\|\widehat{\Phi} b\|_2 \geq \|\widehat{\Phi} e\|_2 - \left\|\widehat{\Phi}(b-e)\right\|_2 \geq (1-\delta'_{\mathcal{M}}) - (N+1)\left(1+\frac{\delta'_{\mathcal{M}}}{4}\right)\epsilon \geq 1 - \frac{9}{4}\delta'_{\mathcal{M}}.$$

Since the upper and lower bounds coincide, and they are valid for all $b \in U(B-B)$, we arrive at our required conclusion. □

D.3 Synthesis

Finally, it remains to extend the stable embedding from $U(B-B)$ to $U(\mathcal{M}-\mathcal{M})$. From Lemma D.1.1, for any $u \in U(\mathcal{M}-\mathcal{M})$, we can find a $b \in U(B-B)$ such that $\|b-u\|_2 \leq \epsilon(T)$ with $\epsilon(T) = 4\sqrt{\frac{T}{\tau}}$. Using Lemma D.2.1 (with ρ fixed and $\delta'_{\mathcal{M}} \leq \frac{4}{9}$ to be defined later), triangle inequalities, and the fact that $\|\Phi\|_2 \leq \left(\frac{N}{S}+1\right)\left(1+\frac{\delta'_{\mathcal{M}}}{4}\right)$, we have

$$\|\widehat{\Phi} u\|_2 \leq \|\widehat{\Phi} b\|_2 + \|\Phi\|_2 \|D_\xi \Psi(b-u)\|_2 \leq (1+\frac{9}{4}\delta'_{\mathcal{M}}) + \left(1+\frac{\delta'_{\mathcal{M}}}{4}\right)(N+1)\,\epsilon(T). \quad (89)$$

Set T such that $\epsilon(T) = \frac{\delta'_{\mathcal{M}}}{N+1}$. By using the formula for $\epsilon(T)$, we have that $T = \frac{(\delta'_{\mathcal{M}})^2 \tau}{16(N+1)^2}$. It is easy to check that $T \leq \tau$, which fulfills the condition of Lemma D.1.1. Plugging this choice of $\epsilon(T)$ into (89), we get

$$\|\widehat{\Phi} u\|_2 \leq (1+\frac{9}{4}\delta'_{\mathcal{M}}) + \left(1+\frac{\delta'_{\mathcal{M}}}{4}\right)\delta'_{\mathcal{M}} \leq 1 + \frac{7}{2}\delta'_{\mathcal{M}},$$

where we used the fact that $\delta'_{\mathcal{M}} \leq \frac{4}{9} < 1$. For the lower conditioning bound, we use the same estimates to arrive at

$$\|\widehat{\Phi} u\|_2 \geq \|\widehat{\Phi} b\|_2 - \|\Phi\|_2 \|D_\xi \Psi(b-u)\|_2 \geq (1-\frac{9}{4}\delta'_{\mathcal{M}}) - \left(1+\frac{\delta'_{\mathcal{M}}}{4}\right)\delta'_{\mathcal{M}} \geq 1 - \frac{7}{2}\delta'_{\mathcal{M}}.$$

Since the upper and lower bounds coincide, we have via the squared and non-squared conditioning bounds

$$\sup_{u \in U(\mathcal{M}-\mathcal{M})} \left| \|\widehat{\Phi} u\|_2^2 - 1 \right| \leq 3 \sup_{u \in U(\mathcal{M}-\mathcal{M})} \left| \|\widehat{\Phi} u\|_2 - 1 \right| \leq \frac{21}{2}\delta'_{\mathcal{M}}.$$

It remains to do some bookkeeping. First, given a predetermined stable manifold embedding conditioning $\delta_{\mathcal{M}} < 1$, set $\delta'_{\mathcal{M}} = \frac{2}{21}\delta_{\mathcal{M}}$. It is clear that this choice of $\delta'_{\mathcal{M}}$ validates the assumption that $\delta'_{\mathcal{M}} \leq \frac{4}{9}$ in Lemma D.2.1, and we have $\sup_{u \in U(\mathcal{M}-\mathcal{M})} \left| \|\widehat{\Phi} u\|_2^2 - 1 \right| \leq \delta_{\mathcal{M}}$ which is what we are trying to prove. Next, according to the JL lemma for RIP operators (Lemma D.2.1), the RIP conditioning for the matrix Φ needs to satisfy $\delta \leq \frac{\delta'_{\mathcal{M}}}{4} = \frac{\delta_{\mathcal{M}}}{42}$. This is the condition for the RIP conditioning in Theorem 4.1.2. Finally, according to the JL lemma for RIP operators (Lemma D.2.1), the RIP order needs to satisfy $S \geq 40 \log\left(\frac{4|E(\delta'_{\mathcal{M}}/(N+1))|}{\rho}\right)$. For this, we need do some work. First, using (87), we have $\left|E\left(\frac{\delta'_{\mathcal{M}}}{N+1}\right)\right| = \left|E\left(\frac{2\delta_{\mathcal{M}}}{21(N+1)}\right)\right| \leq 2|A|^2 \left(1 + \frac{21(N+1)}{\delta_{\mathcal{M}}}\right)^{2D+1}$. Now $|A|$ depends on the geodesic covering resolution T, which was set to be $T = \frac{(\delta'_{\mathcal{M}})^2 \tau}{16(N+1)^2} = \frac{\delta_{\mathcal{M}}^2 \tau}{1764(N+1)^2}$. By Lemma 2.5.2, which gives the geodesic number of a manifold with geodesic regularity R, we have

$$\begin{aligned}
\log(|A|) &\leq \log\left(\frac{\left(\frac{2R}{\sqrt{\pi}}\right)^D \left(\sqrt{D/2+1}\right)^D V}{T^D}\right) \\
&= \log\left(\frac{\left(\frac{3528R}{\sqrt{\pi}}\right)^D \left(\sqrt{D/2+1}\right)^D (N+1)^{2D} V}{\delta_{\mathcal{M}}^{2D} \tau^D}\right) \\
&= D \log\left(\frac{3528R\left(\sqrt{D/2+1}\right)(N+1)^2}{\sqrt{\pi}\delta_{\mathcal{M}}^2 \tau}\right) + \log(V).
\end{aligned}$$

Putting everything together, the order S of the RIP of the matrix Φ must satisfy

$$\begin{aligned}
S \geq\ & 40 \left(2D \log\left(\frac{3528R\left(\sqrt{D/2+1}\right)(N+1)^2}{\sqrt{\pi}\delta_{\mathcal{M}}^2 \tau}\right) + \cdots \right. \\
& \left. \cdots (2D+1) \log\left(1 + \frac{21(N+1)}{\delta_{\mathcal{M}}}\right) + \log\left(\frac{8V^2}{\rho}\right)\right).
\end{aligned}$$

This concludes the proof of Theorem 4.1.2.

APPENDIX E

STABLE TAKENS' EMBEDDING PROOFS

In this appendix, we provide proofs for the stable embedding of dynamical systems' attractors by delay-coordinate maps described in Chapter 5. The proofs involving linear dynamical systems and linear observation functions (Theorems 5.2.1 and 5.2.2) are shown in Section E.1 while the proofs involving nonlinear dynamical systems and general observation functions (Theorems 5.3.1 and 5.3.2) are shown in Section E.2.

E.1 Proof of Stable Takens' Embedding Theorem For Linear Dynamical Systems

Because Theorems 5.2.1 and 5.2.2 are very similar in structure, we will essentially lay out the proof approach for both of them together in this section and then separately establish the necessary details for each result. Before proceeding with the specific proofs, we will introduce some notation and preliminary results that will be useful.

E.1.1 Notation and Preliminaries

E.1.1.1 Frame Theory

Drawing on some terminology from the field of *frame theory*, we say that a sequence of vectors $\{g_i\}_{i=1}^M$ in $\mathbb{C}^K$, $M \geq K$, forms a *frame* [38] for $\mathbb{C}^K$ if there exists two real constants $0 < B_1 \leq B_2 < \infty$ such that for all $\alpha \in \mathbb{C}^K$, $B_1\|\alpha\|_2^2 \leq \sum_{i=1}^M |\langle g_i, \alpha\rangle|^2 = \|G\alpha\|_2^2 \leq B_2\|\alpha\|_2^2$, where $G^H = (g_1 \mid g_2 \mid \cdots \mid g_M) \in \mathbb{C}^{K\times M}$, the concatenation of the $\{g_i\}_{i=1}^M$, is called the *frame analysis operator* and B_1, B_2 are called the *frame bounds*. The frame bounds can be defined as $B_1 = \lambda_{\min}$ and $B_2 = \lambda_{\max}$, where $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum and maximum eigenvalues of $G^HG \in \mathbb{C}^{K\times K}$.

E.1.1.2 Linear Delay Coordinate Maps

Because the attractor $\mathcal{M}$ is contained in the span of the columns of V, for any $x, y \in \mathcal{M}$ we can write $x = V\alpha_x$ and $y = V\alpha_y$ for some complex coefficients $\alpha_x, \alpha_y \in \mathbb{C}^{2d}$. Using F to denote the delay coordinate map for a linear system with flow matrix ψ and observation function h as described in (30), the k-th row (for $k = 1, \cdots, M$) of the vector $F(x) - F(y)$ can be written as

$$\begin{aligned} h^T\left(\psi^{k-1}(x-y)\right) &= h^T\left(\psi^{k-1}V(\alpha_x - \alpha_y)\right) = h^T\left(VD^{k-1}(\alpha_x - \alpha_y)\right) \\ &= \langle g_k, \alpha_x - \alpha_y\rangle, \end{aligned}$$

where

$$\begin{aligned} g_k^H &= h^T V D^{k-1} \\ &= \left[(v_1^T h)e^{-j(k-1)\theta_1 T_s}, (v_1^H h)e^{j(k-1)\theta_1 T_s}, \ldots, (v_d^T h)e^{-j(k-1)\theta_d T_s}, (v_d^H h)e^{j(k-1)\theta_d T_s}\right] \end{aligned} \tag{90}$$

and D is the diagonal matrix comprised of $\mathcal{A}_\psi$-eigenvalues. Thus, we have:

$$\|F(x) - F(y)\|_2^2 = \sum_{k=1}^{M} |\langle g_k, (\alpha_x - \alpha_y)\rangle|^2 = \|G(\alpha_x - \alpha_y)\|_2^2,$$

where $G \in \mathbb{C}^{M\times 2d}$ is the concatenation of $\{g_k\}$ as described above. In this following, G is fixed to be this matrix given here.

E.1.1.3 Eigenvalue Bounds

It will be important in the following proofs to determine bounds on the extreme eigenvalues of the matrix $G^H G$. To that end, we first introduce the well-known Gershgorin Circle Theorem, which we state here for notational convenience:

Theorem E.1.1 (Gershgorin Circle Theorem [98])**.** *The eigenvalues of a $K \times K$ matrix A all lie in the union of the Gershgorin disks of A. The Gershgorin disk $\mathcal{D}_i$ for $i = 1, \cdots, K$, is defined as $\mathcal{D}_i = \{x \in \mathbb{C} \ : \ |x - \mathcal{C}_i| \leq \widetilde{r}_i\}$, where $\widetilde{r}_i := \sum_{j=1,\, j\neq i}^{K} |(A)_{i,j}|$ is*

the radius, and $\mathcal{C}_i := (A)_{i,i}$ *is the center of the* i*-th disk. Thus* $\lambda(A) \subset \bigcup_{i=1}^{K} \mathcal{D}_i$*, where* $\lambda(A) = \{\lambda_1, \cdots, \lambda_K\}$*, and* $\{\lambda_i\}$ *are the eigenvalues of* A*.*

To apply the Gershgorin Circle Theorem to obtain the extrema eigenvalues of $G^H G$, we introduce the following useful lemma that gives values for centers $\mathcal{C}_i$ and radii $\widetilde{r}_i$ of the Gershgorin disks $\mathcal{D}_i$ of $G^H G$.

Lemma E.1.1. *For* $i = 1, \cdots, d$*, the centers of the Gershgorin disks of* $G^H G$ *are*

$$\mathcal{C}_{2i-1} = \mathcal{C}_{2i} = |v_i^H h|^2 M,$$

while their radii are

$$\begin{aligned}\widetilde{r}_{2i-1} = \widetilde{r}_{2i} \;\; = \;\; & |v_i^H h|^2 \left| \frac{\sin(M\theta_i T_s)}{\sin(\theta_i T_s)} \right| + \sum_{p=1,\, p\neq i}^{d} |v_i^H h||v_p^H h| \left| \frac{\sin\left(M(\theta_i - \theta_p)T_s/2\right)}{\sin\left((\theta_i - \theta_p)T_s/2\right)} \right| + \cdots \\ & \cdots + \sum_{p=1,\, p\neq i}^{d} |v_i^H h||v_p^H h| \left| \frac{\sin\left(M(\theta_i + \theta_p)T_s/2\right)}{\sin\left((\theta_i + \theta_p)T_s/2\right)} \right|.\end{aligned}$$

Proof. We can write $G^H G = \sum_{k=1}^{M} g_k g_k^H$, where we recall that g_k is defined as in (90). Thus the (p, q) entry of $G^H G$ can be expressed as: $(G^H G)_{p,q} = \sum_{k=1}^{M} g_k(p) g_k(q)^*$, where $g_k(p)$ denotes the p-th entry of the vector g_k. As such, the formation of $G^H G$ involves the calculation of sum of complex trigonometric polynomials due to the complex exponentials $\left(\{e^{\pm j(k-1)\theta_p T_s}\}\right)$ appearing in the terms of each g_k. A few separate cases need to be considered because of the differences in the even $(2p)$ and odd $(2p-1)$ numbered rows of $G^H G$ for all p. We first consider the even numbered rows. The diagonal terms actually have a fairly simple form: $(G^H G)_{2p,2p} = \sum_{k=1}^{M} g_k(2p) g_k(2p)^* = \sum_{k=1}^{M} |v_p^H h|^2 = M|v_p^H h|^2$. The adjacent term to the left is given by: $(G^H G)_{2p,2p-1} = \sum_{k=0}^{M-1} \left((v_p^T h) e^{-jk\theta_p T_s}\right)^2 = (v_p^T h)^2 \sum_{k=0}^{M-1} \left(e^{-j2\theta_p T_s}\right)^k = (v_p^T h)^2 \frac{\sin(M\theta_p T_s)}{\sin(\theta_p T_s)} e^{-j(M-1)\theta_p T_s}$, where the last expression follows from the standard formula for a finite geometric sum, pulling out common exponential factors, and using Euler's formula. The other cross terms for all $p, q \in \{1, \ldots, d\}$ such that $p \neq q$ can

be derived similarly as:

$$\begin{aligned}
(G^HG)_{2p,2q} &= (v_p^T h)(v_q^H h)\sum_{k=0}^{M-1}\left(e^{-j2\left(\frac{\theta_p-\theta_q}{2}\right)T_s}\right)^k \\
&= (v_p^T h)(v_q^H h)\frac{\sin\left(M\left(\frac{\theta_p-\theta_q}{2}\right)T_s\right)}{\sin\left(\left(\frac{\theta_p-\theta_q}{2}\right)T_s\right)}e^{-j(M-1)\left(\frac{\theta_p-\theta_q}{2}\right)T_s}, \\
(G^HG)_{2p,2q-1} &= (v_p^T h)(v_q^T h)\sum_{k=0}^{M-1}\left(e^{-j2\left(\frac{\theta_p+\theta_q}{2}\right)T_s}\right)^k \\
&= (v_p^T h)(v_q^T h)\frac{\sin\left(M\left(\frac{\theta_p+\theta_q}{2}\right)T_s\right)}{\sin\left(\left(\frac{\theta_p+\theta_q}{2}\right)T_s\right)}e^{-j(M-1)\left(\frac{\theta_p+\theta_q}{2}\right)T_s}.
\end{aligned}$$

The relevant quantities for the odd numbered rows are given similarly as

$$\begin{aligned}
(G^HG)_{2p-1,2p-1} &= (G^HG)_{2p,2p} = M|v_p^H h|^2, \\
(G^HG)_{2p-1,2p} &= (G^HG)^*_{2p,2p-1} = (v_p^H h)^2\frac{\sin(M\theta_p T_s)}{\sin(\theta_p T_s)}e^{j(M-1)\theta_p T_s}, \\
(G^HG)_{2p-1,2q} &= (G^HG)^*_{2q,2p-1} = (v_q^H h)(v_p^H h)\frac{\sin\left(M(\theta_q+\theta_p)T_s/2\right)}{\sin\left((\theta_q+\theta_p)T_s/2\right)}e^{j(M-1)\left(\frac{\theta_q+\theta_p}{2}\right)T_s}, \\
(G^HG)_{2p-1,2q-1} &= (v_p^H h)(v_q^T h)\frac{\sin\left(M(\theta_p-\theta_q)T_s/2\right)}{\sin\left((\theta_p-\theta_q)T_s/2\right)}e^{j(M-1)\left(\frac{\theta_p-\theta_q}{2}\right)T_s}.
\end{aligned}$$

Finally we note that many of the above complex quantities only differ in their phase because of symmetry in the summations, making their magnitudes equal when calculating the radii of the Gershgorin disks. The expressions for $\mathcal{C}_i$ and $\widetilde{r}_i$ in the lemma are obtained simply by applying the notation of the Gershgorin Circle Theorem to the calculated magnitudes of the entries of G^HG. □

E.1.2 General Proof Approach

Using the preliminaries above, we can now sketch out the general approach for the proof of both theorems below. Essentially, the theorems result from using (or establishing) the following three facts:

1. If $G^HG \in \mathbb{C}^{2d\times 2d}$ is established to be full rank, then $\{g_k\}_{k=1}^M$ form a frame in $\mathbb{C}^{2d}$. Thus, there exists $0 < B_1 \leq B_2 < \infty$ such that $B_1 \leq \frac{\|F(x)-F(y)\|_2^2}{\|\alpha_x-\alpha_y\|_2^2} \leq B_2$ holds

for all distinct pairs of points $x, y \in \mathcal{M}$. In particular, to establish conditioning guarantees, we can let B_1 and B_2 be the smallest and largest eigenvalues of $G^H G$ (respectively) and determine bounds on those important quantities.

2. Next, we use the fact that $\|x-y\|_2^2 = (\alpha_x - \alpha_y)^H V^H V (\alpha_x - \alpha_y)$ to get $A_1 \leq \frac{\|x-y\|_2^2}{\|\alpha_x - \alpha_y\|_2^2} \leq A_2$, where A_1 and A_2 are the smallest and largest eigenvalues of $V^H V \in \mathbb{C}^{2d \times 2d}$ respectively. By the definition of V we know that $V^H V$ is well-defined and full rank, meaning that $0 < A_1 \leq A_2 < \infty$.

3. Putting the 2 previous steps together, we get $0 < \frac{B_1}{A_2} \leq \frac{\|F(x)-F(y)\|_2^2}{\|x-y\|_2^2} \leq \frac{B_2}{A_1} < \infty$, where the bounds $\frac{B_1}{A_2}$ and $\frac{B_2}{A_1}$ can be manipulated to get the scaling constant C and conditioning δ in (31). Specifically, we can set $C = \frac{1}{2}\left(\frac{B_1}{A_2} + \frac{B_2}{A_1}\right)$ and $\delta = 1 - \frac{B_1}{CA_2}$.

E.1.3 Proof of Theorem 5.2.1

Proof. For Theorem 5.2.1, we follow the three steps detailed in Appendix E.1.2, where we only need to show that $G^H G$ is indeed full rank given the conditions of the theorem. Consider first the case when $M = 2d$, where showing $G^H G$ is full rank is equivalent to showing $\det(G^H G) = \det(G)^2 > 0$. The matrix G can be expressed in terms of a product of a Vandermonde matrix and a diagonal matrix:

$$G = \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ e^{-j\theta_1 T_s} & e^{j\theta_1 T_s} & \cdots & e^{-j\theta_d T_s} & e^{j\theta_d T_s} \\ \vdots & \vdots & & \vdots & \vdots \\ e^{-j2d\theta_1 T_s} & e^{j2d\theta_1 T_s} & \cdots & e^{-j2d\theta_d T_s} & e^{j2d\theta_d T_s} \end{pmatrix} \begin{pmatrix} v_1^T h & & & & (0) \\ & v_1^H h & & & \\ & & \ddots & & \\ & & & v_d^T h & \\ (0) & & & & v_d^H h \end{pmatrix} = \widetilde{M}^T \widetilde{H},$$

where $\widetilde{M}$ is the Vandermonde matrix with the $\mathcal{A}_\psi$-eigenvalues as its parameters and $\widetilde{H}$ is a diagonal matrix whose diagonal elements are made up of the projection of h onto the $\mathcal{A}$-eigenvectors. Thus, $\det(G) = \det(\widetilde{M})\det(\widetilde{H})$. One of the conditions of Theorem 5.2.1 ensures that the $\{e^{\pm j\theta_i T_s}\}_{i=1}^d$ are distinct, which implies that the determinant of this square Vandermonde matrix [75, Ch 0] obeys $|\det(\widetilde{M})| > 0$. Also since $v_i^H h \neq 0$ for all $i = 1, \cdots, d$, we also know that $|\det(\widetilde{H})| > 0$. Therefore for

$M = 2d$, $\text{rank}(G^H G) = 2d$. Since adding vectors to a frame does not change the rank of $G^H G$ (i.e., frame bounds cannot be lowered by adding more vectors to the frame), it follows that if $M \geq 2d$ then $\text{rank}(G^H G) = 2d$ and the proof of Theorem 5.2.1 is complete. □

E.1.4 Proof of Theorem 5.2.2

Proof. To prove Theorem 5.2.2, we again follow the three steps detailed in Appendix E.1.2, this time establishing specific guarantees on the frame bounds $B_1(M)$ and $B_2(M)$ appearing in the first step. From Lemma E.1.1, we first observe that for all i we can bound the Gershgorin disk radii by

$$\begin{aligned} \widetilde{r}_{2i-1} = \widetilde{r}_{2i} &\leq \left(|v_i^H h|^2 + \sum_{p=1,\, p\neq i}^{d} |v_i^H h||v_p^H h| + \sum_{p=1,\, p\neq i}^{d} |v_i^H h||v_p^H h| \right) \nu \\ &\leq (2d-1)\kappa_2^2 \|h\|_2^2 \nu. \end{aligned}$$

Noting that $\|h\|_2^2 = \frac{2d}{M}$, we see that for each i, the Gershgorin disks of $G^H G$ satisfy

$$\mathcal{D}_{2i-1} = \mathcal{D}_{2i} \subset \left[|v_i^H h|^2 M - \|h\|_2^2 (2d-1)\nu\kappa_2^2,\ |v_i^H h|^2 M + \|h\|_2^2 (2d-1)\nu\kappa_2^2 \right].$$

Then applying the Gershgorin Circle Theorem, we get

$$\lambda(G^H G) \subset \bigcup_j^{2d} \mathcal{D}_j \subset \left[2d\kappa_1^2 - \frac{2d}{M}(2d-1)\nu\kappa_2^2,\ 2d\kappa_2^2 + \frac{2d}{M}(2d-1)\nu\kappa_2^2 \right].$$

By choosing $B_1(M) = 2d\left(\kappa_1^2 - \frac{(2d-1)\nu\kappa_2^2}{M}\right)$ and $B_2(M) = 2d\left(\kappa_2^2 + \frac{(2d-1)\nu\kappa_2^2}{M}\right)$, and applying step 2 in Section E.1.2, we arrive at

$$\frac{B_1(M)}{A_2} \leq \frac{\|F(x) - F(y)\|_2^2}{\|x - y\|_2^2} \leq \frac{B_2(M)}{A_1} \tag{91}$$

for all distinct pairs of points $x, y \in \mathcal{M}$ and for all M.

Now as $M \to \infty$, $B_1(M) \to 2d\kappa_1^2$ and $B_2(M) \to 2d\kappa_2^2$. Thus in the limit of large M, the lower and upper bounds of the inequality (91) approaches $\frac{2d\kappa_1^2}{A_2}$ and $\frac{2d\kappa_2^2}{A_1}$, respectively. We define the scaling constant C as the average of the asymptotic values

of these lower and upper bounds: $C := \frac{2d}{2}\left(\frac{\kappa_1^2}{A_2} + \frac{\kappa_2^2}{A_1}\right)$. Also define the conditioning number $\delta(M)$ for a given M as the maximum deviation of the lower and upper bounds of (91) from C, normalized by C: $\delta(M) := \max\left\{1 - \frac{B_1(M)}{CA_2}, \frac{B_2(M)}{CA_1} - 1\right\}$. Now $1 - \frac{B_1(M)}{CA_2} = 1 - \frac{2d\left(\kappa_1^2 - (2d-1)\nu\kappa_2^2/M\right)}{(2d/2)\left(\kappa_1^2 + \kappa_2^2(A_2/A_1)\right)} = \frac{A_2\kappa_2^2 - A_1\kappa_1^2 + 2A_1(2d-1)\nu\kappa_2^2/M}{A_2\kappa_2^2 + A_1\kappa_1^2}$, and $\frac{B_2(M)}{CA_1} - 1 = \frac{2d\left(\kappa_2^2 + (2d-1)\nu\kappa_2^2/M\right)}{(2d/2)\left((A_1/A_2)\kappa_1^2 + \kappa_2^2\right)} - 1 = \frac{A_2\kappa_2^2 - A_1\kappa_1^2 + 2A_2(2d-1)\nu\kappa_2^2/M}{A_2\kappa_2^2 + A_1\kappa_1^2}$. Since $A_1 \leq A_2$, we have that $\delta(M) = \frac{B_2(M)}{CA_1} - 1 = \frac{A_2\kappa_2^2 - A_1\kappa_1^2}{A_2\kappa_2^2 + A_1\kappa_1^2} + \frac{2A_2\kappa_2^2}{A_2\kappa_2^2 + A_1\kappa_1^2}\frac{(2d-1)\nu}{M}$. We can then define δ_0 and $\delta_1(M)$ as the first and second term of the sum above. Notice that $\delta(M)$ represents a worst case bound on the deviation from C, as we maximized over upper and lower bounds that may not be the same magnitude (i.e., in general $C(1-\delta(M)) \neq \frac{B_1(M)}{A_2}$).

Finally, we recall that for the embedding conditioning number to be valid, we must have $0 \leq \delta(M) < 1$. The first condition $\delta(M) \geq 0$ is achieved by construction. The upper bound is equivalent to the condition for M required by the theorem statement, thus completing the proof. □

E.2 Proof of Stable Takens' Embedding for Nonlinear Dynamical Systems and Nonlinear Observation Functions

In this section, we will prove Theorem 5.3.2. It is easy to see that Theorem 5.3.1 is then a direct corollary of this theorem.

Let $\mathcal{M}$ be the Riemannian submanifold considered in Theorem 5.3.2. Recall that F_α is a stable embedding of $\bar{g}(\mathcal{M})$ with conditioning δ whenever

$$(1-\delta) \leq \frac{\|F_\alpha(x) - F_\alpha(y)\|_2^2}{\|\bar{g}(x) - \bar{g}(y)\|_2^2} = \frac{\|F_\alpha(x) - F_\alpha(y)\|_2^2}{\|G_x - G_y\|_F^2} \leq (1+\delta)$$

holds for all $\bar{g}(x), \bar{g}(y) \in \bar{g}(\mathcal{M})$ (or for all $x, y \in \mathcal{M}$ since $\bar{g}(\cdot)$ is one-to-one on $\mathcal{M}$). Recall the fact that for some $x \in \mathcal{M}$, we can rewrite $F_\alpha(x) = G_x\alpha$ with α being an i.i.d. Rademacher sequence, and that G_x is the matrix equivalent of the vector $\bar{g}(x) \in \bar{g}(\mathcal{M})$. Without loss of generality, we shall write $G_x \in \bar{g}(\mathcal{M})$. Then, we can

rewrite the statement that F_α is a stable embedding of $\bar{g}(\mathcal{M})$ with conditioning δ as

$$\begin{aligned}\sup_{x,y\in\mathcal{M}}\left|\frac{\|F_\alpha(x)-F_\alpha(y)\|_2^2}{\|G_x-G_y\|_F^2}-1\right| &= \sup_{U(G_x-G_y)\in U(\bar{g}(\mathcal{M})-\bar{g}(\mathcal{M}))}\left|\|U(G_x-G_y)\alpha\|_2^2-1\right| \\ &\leq \delta, \end{aligned} \tag{92}$$

where from our notations, $U(G_x-G_y) := \frac{G_x-G_y}{\|G_x-G_y\|_F}$. Thus, showing stable embedding is equivalent to showing that the supremum of the random variable $\left|\|U(G_x-G_y)\alpha\|_2^2-1\right|$ over the set $U(\bar{g}(\mathcal{M})-\bar{g}(\mathcal{M}))$ is bounded by δ. For notational simplicity, we denote $\bar{U} := U(\bar{g}(\mathcal{M})-\bar{g}(\mathcal{M}))$.

To get a bound on the supremum of the random variable $\left|\|U(G_x-G_y)\alpha\|_2^2-1\right|$, recall that we can also write $F_\alpha(x)-F_\alpha(y)$ as (see Section 5.3.4)

$$F_\alpha(x)-F_\alpha(y) = (G_x-G_y)\alpha = \begin{pmatrix}\alpha^T & & (0)\\ & \ddots & \\ (0) & & \alpha^T\end{pmatrix}(\bar{g}(x)-\bar{g}(y)) =: \Phi(\bar{g}(x)-\bar{g}(y)), \tag{93}$$

where $\Phi\in\mathbb{R}^{M\times MP}$ is a RBD matrix (see Section 3.1.2 and [107]). The following lemma gives a concentration of measure bound on the RBD matrix Φ as found in Section 3.1.2 and [84, Proposition 5.2]:

Lemma E.2.1. *Suppose Φ is a $M\times MP$ RBD matrix appearing in (93) where $\alpha\in\mathbb{R}^P$ is either an i.i.d. Rademacher or standard Gaussian sequence. Then, for any $\bar{g}(x),\bar{g}(y)\in\bar{g}(\mathcal{M})$,*

$$\mathbb{P}\left\{\left|\frac{\|\Phi(\bar{g}(x)-\bar{g}(y))\|_2^2}{\|\bar{g}(x)-\bar{g}(y)\|_2^2}-1\right|>\epsilon\right\} \leq 2\exp\left(-C(\epsilon)r(G_x-G_y)\right), \tag{94}$$

where we recall that $r(G_x-G_y) := \frac{\|G_x-G_y\|_F^2}{\|G_x-G_y\|_2^2}$ is the soft-rank of G_x-G_y, and $C(\epsilon) := \min\left\{\frac{\epsilon}{16},\left(\frac{\epsilon}{16}\right)^2\right\}$ if α is a Gaussian sequence and $C(\epsilon) := \frac{1}{64}\min\left\{\frac{96\epsilon}{130},\epsilon^2\right\}$ if α is a Rademacher sequence.

Proof. We will start with the case when α is an i.i.d. standard Gaussian sequence. From Theorem 3.1.2, we have that

$$\mathbb{P}\left\{\left|\frac{\|\Phi(\bar{g}(x)-\bar{g}(y))\|_2^2}{\|\bar{g}(x)-\bar{g}(y)\|_2^2}-1\right|>\epsilon\right\} \leq 2\exp\left(-\min\left\{\left(\frac{\epsilon}{16}\right)^2\frac{\|\lambda\|_1^2}{\|\lambda\|_2^2},\ \frac{\epsilon}{16}\frac{\|\lambda\|_1}{\|\lambda\|_\infty}\right\}\right),$$

for any $\bar{g}(x), \bar{g}(y) \in \mathbb{R}^{MP}$, and where $\lambda \in \mathbb{R}^{MP}$ are the singular values of $(G_x - G_y)^H(G_x - G_y)$. Then, using the inequality $\|\lambda\|_2^2 \leq \|\lambda\|_1\|\lambda\|_\infty$ and the fact that $\frac{\|\lambda\|_1}{\|\lambda\|_\infty} = \frac{\|G_x - G_y\|_F^2}{\|G_x - G_y\|_2^2} = r(G_x - G_y)$ yields (94).

Moving on to the case when α is an i.i.d. Rademacher sequence, [84, Proposition 5.2] tells us that for any $\bar{g}(x), \bar{g}(y) \in \mathbb{R}^{MP}$,

$$\begin{aligned} & \mathbb{P}\Big\{\big|\|\Phi(\bar{g}(x) - \bar{g}(y))\|_2^2 - \|\bar{g}(x) - \bar{g}(y)\|_2^2\big| > t\Big\} \\ \leq \; & 2\exp\left(-\frac{1}{64}\min\left\{\frac{\frac{96}{65}t}{\|X\|_2}, \frac{t^2}{\|X\|_F^2}\right\}\right), \end{aligned} \tag{95}$$

where $X := H((G_x - G_y)^H(G_x - G_y))$ with $H(A)$ being the hollow function of a square matrix A that returns the matrix A with its diagonal set to 0. First, it is easy to see that

$$\|H((G_x - G_y)^H(G_x - G_y))\|_F^2 \leq \|(G_x - G_y)^H(G_x - G_y)\|_F^2 = \|\lambda\|_2^2,$$

where λ is again the vector of singular values of $(G_x - G_y)^H(G_x - G_y)$. Next, we have

$$\|H((G_x - G_y)^H(G_x - G_y))\|_2 \leq \|(G_x - G_y)^H(G_x - G_y)\|_2 + \|D((G_x - G_y)^H(G_x - G_y))\|_2,$$

where $D(A)$ is a diagonal function that returns the matrix A but with zeros on all entries except on the diagonal. Now,

$$\|D((G_x - G_y)^H(G_x - G_y))\|_2 \leq \|(G_x - G_y)^H(G_x - G_y)\|_{\max},$$

where $\|A\|_{\max}$ returns the maximum absolute value of the entries of A, and

$$\|(G_x - G_y)^H(G_x - G_y)\|_{\max} \leq \|(G_x - G_y)^H(G_x - G_y)\|_2.$$

Therefore, we have

$$\|H((G_x - G_y)^H(G_x - G_y))\|_2 \leq 2\|(G_x - G_y)^H(G_x - G_y)\|_2 = 2\|\lambda\|_\infty.$$

Putting these bounds on $\|X\|_2$ and $\|X\|_F^2$ into (95) together with the substitution $t = \epsilon\|\bar{g}(x) - \bar{g}(y)\|_2^2$, we have

$$\mathbb{P}\left\{\left\{\left|\frac{\|\Phi(\bar{g}(x) - \bar{g}(y))\|_2^2}{\|\bar{g}(x) - \bar{g}(y)\|_2^2} - 1\right| > \epsilon\right\} \leq 2\exp\left(-\frac{1}{64}\min\left\{\frac{\frac{96}{130}\epsilon\|\lambda\|_1}{\|\lambda\|_\infty}, \frac{\epsilon^2\|\lambda\|_1^2}{\|\lambda\|_2^2}\right\}\right),\right.$$

where we used the fact that $\|\bar{g}(x) - \bar{g}(y)\|_2^2 = \|G_x - G_y\|_F^2 = \|\lambda\|_1$. Again, using the inequality $\|\lambda\|_2^2 \leq \|\lambda\|_1 \|\lambda\|_\infty$ and the fact that $\frac{\|\lambda\|_1}{\|\lambda\|_\infty} = \frac{\|G_x - G_y\|_F^2}{\|G_x - G_y\|_2^2} = r(G_x - G_y)$ yields (94). □

In what follows, we will suppose that α is an i.i.d. Rademacher sequence (and the case when α is an i.i.d. Gaussian sequence is discussed in Section 5.3.4). Noting that $U(G_x - G_y)\alpha = \frac{\Phi(\bar{g}(x) - \bar{g}(y))}{\|\bar{g}(x) - \bar{g}(y)\|_2^2}$, we will use the following concentration of measure inequality which holds for any $U(G_x - G_y) \in \bar{U}$:

$$\mathbb{P}\left\{\left|\|U(G_x - G_y)\alpha\|_2^2 - 1\right| > \epsilon\right\} \leq 2\exp\left(-C(\epsilon) r(G_x - G_y)\right), \tag{96}$$

where $C(\epsilon) := \frac{1}{64}\min\left\{\frac{96\epsilon}{130},\ \epsilon^2\right\}$.

For a fixed stable embedding conditioning δ, we want to show that the following failure probability can be bounded by some predetermined $0 < \rho < 1$:

$$\mathbb{P}\left\{\sup_{U(G_x - G_y) \in \bar{U}} \left|\|U(G_x - G_y)\alpha\|_2^2 - 1\right| > \delta\right\} \leq \rho.$$

To continue our analysis, we will first cover the space of chords $\bar{U}$ with a finite set $\bar{K}$ such that every point in $\bar{U}$ is close to a point in $\bar{K}$. This is the subject of Lemma E.2.2 below. Then, using the union bound together with the concentration of measure bound (96), we can bound

$$\mathbb{P}\left\{\sup_{U(G_x - G_y) \in \bar{K}} \left|\|U(G_x - G_y)\alpha\|_2^2 - 1\right| > \delta\right\}.$$

Lastly, we will use simple geometric arguments to extend the supremum from the set $\bar{K}$ to the set of interest $\bar{U}$.

First, the following lemma shows that a finite cover of $\bar{U}$ exists:

Lemma E.2.2. *Pick any $\eta > 0$. There exists a finite set $\bar{K}$ such that $\bar{K}$ is a $(2\eta, \|\cdot\|_2)$-cover of $\bar{U}$, $\bar{K} \subset \bar{U}$, and*

$$|\bar{K}| \leq 2V_{\bar{g}(\mathcal{M})}^2 \left(\frac{32}{\sqrt{\pi}} \frac{R_{\bar{g}(\mathcal{M})}\sqrt{D_{\mathcal{M}}}}{\tau_{\bar{g}(\mathcal{M})}} \frac{1}{\eta^2}\right)^{2D_{\mathcal{M}}} \left(1 + \frac{4}{\eta}\right)^{2D_{\mathcal{M}}+1},$$

where $D_{\mathcal{M}}$, $R_{\bar{g}(\mathcal{M})}$, $V_{\bar{g}(\mathcal{M})}$, and $\frac{1}{\tau_{\bar{g}(\mathcal{M})}}$ are the dimension, geodesic regularity, volume, and uniform upper bound on the second fundamental form of $\bar{g}(\mathcal{M})$ respectively.

Proof. The proof of this lemma follows from Lemma D.1.1 in Section D.1 and the discussion on $E(\epsilon)$ in Section D.2. To be precise, let $A = A(T) := \mathcal{C}\left(\bar{g}(\mathcal{M}), d_{\bar{g}(\mathcal{M})}, T\right)$ for some $T \leq \tau_{\bar{g}(\mathcal{M})}$ be the $(T, d_{\bar{g}(\mathcal{M})})$-cover of $\bar{g}(\mathcal{M})$ of minimum cardinality. Define a generalized covering set B of the manifold $\bar{g}(\mathcal{M})$ as

$$B = B(T) = \bigcup_{a \in A} \{a + \mathcal{T}_a \bar{g}(\mathcal{M})(T)\},$$

where $\mathcal{T}_a \bar{g}(\mathcal{M})(T) := \{u \in \mathcal{T}_a \bar{g}(\mathcal{M}) \mid \|u\|_2 \leq T\}$ refers to all tangent vectors of $\bar{g}(\mathcal{M})$ at the point a whose lengths are less than T. Then, Lemma D.1.1 in Section D.1 says that $U(B - B)$ is a $\left(4\sqrt{\frac{T}{\tau}}, \|\cdot\|_2\right)$-cover of $\bar{U}$.

Now from Section D.2, $U(B-B)$ can be divided into two sets of elements, namely:

1. $B_1 := U\left(\bigcup_{a \in A} \{\mathcal{T}_a \bar{g}(\mathcal{M})(T) - \mathcal{T}_a \bar{g}(\mathcal{M})(T)\}\right) = U\left(\bigcup_{a \in A} \mathcal{T}_a \bar{g}(\mathcal{M})\right)$, and
2. $B_2 := U\left(\bigcup_{a_1, a_2 \in A, a_1 \neq a_2} \{(a_1 - a_2) + (\mathcal{T}_{a_1} \bar{g}(\mathcal{M})(T) - \mathcal{T}_{a_2} \bar{g}(\mathcal{M})(T))\}\right)$.

The set B_1 is comprised of $|A|$ $D_{\mathcal{M}}$-dimensional unit spheres. We know that each unit sphere can be $(\epsilon, \|\cdot\|_2)$-covered by at most $(1 + \frac{2}{\epsilon})^{D_{\mathcal{M}}}$ points, thus $|\mathcal{C}(B_1, \|\cdot\|_2, \epsilon)| \leq |A|(1 + \frac{2}{\epsilon})^{D_{\mathcal{M}}}$. The set B_2 is the projection onto the unit sphere (in $\mathbb{R}^{MP}$) of not more than $|A|^2$ subsets of affine planes where each affine plane is contained in a linear subspace of dimension $2D_{\mathcal{M}} + 1$. Thus, $|\mathcal{C}(B_2, \|\cdot\|_2, \epsilon)| \leq |A|^2(1 + \frac{2}{\epsilon})^{2D_{\mathcal{M}}+1}$. Define the collection of points $E = E(\epsilon) := \mathcal{C}(B_1, \|\cdot\|_2, \epsilon) \cup \mathcal{C}(B_2, \|\cdot\|_2, \epsilon)$ where

$$|E(\epsilon)| \leq |A| \left(1 + \frac{2}{\epsilon}\right)^{D_{\mathcal{M}}} + |A|^2 \left(1 + \frac{2}{\epsilon}\right)^{2D_{\mathcal{M}}+1} \leq 2|A|^2 \left(1 + \frac{2}{\epsilon}\right)^{2D_{\mathcal{M}}+1}.$$

By construction, for any $b \in U(B-B)$, we can find an $e \in E(\epsilon)$ such that $\|b-e\|_2 \leq \epsilon$. Set $\epsilon = \frac{\eta}{2}$ and $T = \frac{\eta^2 \tau_{\bar{g}(\mathcal{M})}}{64}$. With this choice of ϵ and T, for every $u \in \bar{U}$, we can find a $b \in U(B - B)$ and $e \in E(\epsilon)$ such that $\|u - e\|_2 \leq \|u - b\|_2 + \|b - e\|_2 \leq$

$\frac{\eta}{2} + \frac{\eta}{2} = \eta$. To bound $|A|$, Lemma 2.5.2 of Section 2.5.1 tells us that $|A(T)| \leq \frac{\left(\frac{2R_{\bar{g}(\mathcal{M})}}{\sqrt{\pi}}\right)^{D_{\mathcal{M}}}\left(\sqrt{\frac{D_{\mathcal{M}}}{2}}+1\right)^{D_{\mathcal{M}}} V_{\bar{g}(\mathcal{M})}}{T^{D_{\mathcal{M}}}}$.

However, the points that make up E do not all lie on the set $\bar{U}$ as required by the lemma. To rectify this, for every $e \in E$, define $k_e := \arg\min_{k \in \bar{U}} \|k - e\|_2$. It is easy to check that $\|k_e - e\|_2 \leq \eta$ since E is a $(\eta, |\cdot\|_2)$-cover of $\bar{U}$. Next, define $\bar{K} := \{k_e \mid e \in E\} \subset \bar{U}$ as the set of all k_e (after pruning away repetitions). By construction $|\bar{K}| \leq |E|$ and (for $D_{\mathcal{M}} \geq 2$)

$$|E| \leq 2V^2_{\bar{g}(\mathcal{M})} \left(\frac{128\sqrt{2}}{\sqrt{\pi}} \frac{R_{\bar{g}(\mathcal{M})}\sqrt{D_{\mathcal{M}}}}{\tau_{\bar{g}(\mathcal{M})}} \frac{1}{\eta^2}\right)^{2D_{\mathcal{M}}} \left(1 + \frac{4}{\eta}\right)^{2D_{\mathcal{M}}+1}.$$

We claim that $\bar{K}$ is a $(2\eta, \|\cdot\|_2)$-cover of $\bar{U}$. To prove this, first pick any $u \in \bar{U}$. Then, there exists an $e \in E$ such that $\|u - e\|_2 \leq \eta$. Now, there also exists a $k \in \bar{K}$ such that $\|e - k\|_2 \leq \eta$. Thus, $\|u - k\|_2 \leq \|u - e\|_2 + \|e - k\|_2 \leq 2\eta$ which concludes the proof. □

With this set $\bar{K}$ and using the union bound, we have

$$\begin{aligned}
&\mathbb{P}\left\{\max_{U(\bar{g}(x)-\bar{g}(y)) \in \bar{K}} \left|\|U\left(G_x - G_y\right)\alpha\|_2^2 - 1\right| > \delta\right\} \\
\leq\ & 2|\bar{K}| \exp\left(-C(\delta) \min_{U(G_x - G_y) \in \bar{K}} r(G_x - G_y)\right) \\
\leq\ & 2|\bar{K}| \exp\left(-C(\delta) \inf_{U(G_x - G_y) \in \bar{U}} r(G_x - G_y)\right).
\end{aligned}$$

To complete the proof, we require a simple geometric argument to extend our failure probability from the covering set $\bar{K}$ to the whole of $\bar{U}$. For this, suppose that we are under the event

$$F_1(\delta) := \max_{U(G_x - G_y) \in \bar{K}} \left|\|U\left(G_x - G_y\right)\alpha\|_2^2 - 1\right| \leq \delta.$$

Denoting F^c as the complement of some event F, we have shown that

$$\mathbb{P}\{F_1^c(\delta)\} \leq 2|\bar{K}| \exp\left(-C(\delta) \inf_{U(G_x - G_y) \in \bar{U}} r(G_x - G_y)\right).$$

With the relation between a squared and non-squared conditioning (3), we can deduce that for any $k \in \bar{K}$, we have $|\|\Phi k\|_2 - 1| \leq \delta$ under event $F_1(\delta)$ (non-squared conditioning). Now, pick any $U(G_x - G_y) \in \bar{U}$. By definition of $\bar{K}$, we can find a $k \in \bar{K}$ such that $\|U\left(\bar{g}\left(x\right) - \bar{g}\left(y\right)\right) - k\|_2 \leq 2\eta$. Therefore, for this choice of $U(G_x - G_y) \in \bar{U}$,

$$\begin{aligned}\|U\left(G_x - G_y\right)\alpha\|_2 = \|\Phi U\left(\bar{g}\left(x\right) - \bar{g}\left(y\right)\right)\|_2 &\leq \|\Phi(U\left(\bar{g}\left(x\right) - \bar{g}\left(y\right)\right) - k)\|_2 + \|\Phi k\|_2 \\ &\leq 2\|\Phi\|_2\eta + (1+\delta) = 1 + (\delta + 2\|\Phi\|_2\eta)\,.\end{aligned}$$

Similarly, we have the lower bound

$$\begin{aligned}\|U\left(G_x - G_y\right)\alpha\|_2 = \|\Phi U\left(\bar{g}\left(x\right) - \bar{g}\left(y\right)\right)\|_2 &\geq \|\Phi k\|_2 - \|\Phi(U\left(\bar{g}\left(x\right) - \bar{g}\left(y\right)\right) - k)\|_2 \\ &\geq (1-\delta) - 2\|\Phi\|_2\eta = 1 - (\delta + 2\|\Phi\|_2\eta)\,.\end{aligned}$$

Combining these bounds, we have that

$$|\|U\left(G_x - G_y\right)\alpha\|_2 - 1| \leq \delta + 2\|\Phi\|_2\eta. \tag{97}$$

Now, a crude bound gives $\|\Phi\|_2^2 \leq \|\Phi\|_F^2 = PM$ since α is a Rademacher sequence. Thus, under the event $F_1(\delta)$ and choosing $\eta = \frac{\delta}{2PM}$, we have (non-square conditioning)

$$|\|U\left(G_x - G_y\right)\alpha\|_2 - 1| \leq 2\delta,$$

for any $U\left(G_x - G_y\right) \in \bar{U}$. Then, applying the relation between squared and non-squared conditioning (3) gives

$$\left|\|U\left(G_x - G_y\right)\alpha\|_2^2 - 1\right| \leq 6\delta.$$

Piecing everything together, we have that

$$\begin{aligned}&\mathbb{P}\left\{\sup_{U(G_x - G_y)\in\bar{U}}\left|\|U\left(G_x - G_y\right)\alpha\|_2^2 - 1\right| > \delta\right\} \\ \leq\ &\mathbb{P}\left\{\max_{U(G_x - G_y)\in\bar{K}}\left|\|U\left(G_x - G_y\right)\alpha\|_2^2 - 1\right| > \frac{\delta}{6}\right\} \\ \leq\ &2|\bar{K}|\exp\left(-C\left(\frac{\delta}{6}\right)\inf_{U(G_x - G_y)\in\bar{U}} r(G_x - G_y)^2\right).\end{aligned} \tag{98}$$

To bound the right-hand-side of (98) with a fixed failure probability ρ implies that we need

$$\inf_{U(G_x-G_y)\in\bar{U}} r(G_x-G_y)^2 \geq \frac{\log(2|\bar{K}|)+\log\left(\frac{1}{\rho}\right)}{C\left(\frac{\delta}{6}\right)}.$$

Since $|\bar{K}| \leq 2V^2_{\bar{g}(\mathcal{M})}\left(\frac{128\sqrt{2}}{\sqrt{\pi}}\frac{R_{\bar{g}(\mathcal{M})}\sqrt{D_{\mathcal{M}}}}{\tau_{\bar{g}(\mathcal{M})}}\frac{1}{\eta^2}\right)^{2D_{\mathcal{M}}}\left(1+\frac{4}{\eta}\right)^{2D_{\mathcal{M}}+1}$ with $\eta=\frac{\delta}{2PM}$, we arrive at the requirement that

$$\begin{aligned}\inf_{U(G_x-G_y)\in\bar{U}} r(G_x-G_y)^2 \;\geq\; & \frac{1}{C\left(\frac{\delta}{6}\right)}\left[2D_{\mathcal{M}}\log\left(\frac{128\sqrt{2}}{\sqrt{\pi}}\frac{R_{\bar{g}(\mathcal{M})}\sqrt{D_{\mathcal{M}}}}{\tau_{\bar{g}(\mathcal{M})}}\right)+4D_{\mathcal{M}}\log\left(\frac{2PM}{\delta}\right)\right.\\ & \left.+(2D_{\mathcal{M}}+1)\log\left(1+\frac{8PM}{\delta}\right)+\log\left(\frac{4V^2_{\bar{g}(\mathcal{M})}}{\rho}\right)\right].\end{aligned}$$

Now, because $C\left(\frac{\delta}{6}\right)=\frac{1}{64}\min\left\{\frac{96\delta}{390},\left(\frac{\delta}{6}\right)^2\right\}=\frac{1}{64}\cdot\frac{\delta^2}{36}=\frac{\delta^2}{2304}$ since $\frac{96\delta}{390}>\frac{\delta}{6}>\left(\frac{\delta}{6}\right)^2$ for $\delta<1$, we obtain the bound in the theorem.

APPENDIX F

MEMORY IN NEURAL NETWORKS PROOFS

In this appendix, we prove the statements that appeared in Chapter 6.

F.1 Proof of Basic RIP Statement

In this section, we show that the matrix $\Phi = U\widetilde{Z}F$ satisfies the RIP under the conditions stated in (44) in Section 6.3. Specifically, we want to prove the following theorem:

Theorem F.1.1. *Suppose $N \geq M$, $N \geq S$ and $N \geq O(1)$. Let $z = \frac{1}{\sqrt{M}} U \mathbf{1}_M$ where the eigenvector matrix U is unitary, so that $\widetilde{Z} = \operatorname{diag}(U^{-1}z) = \frac{1}{\sqrt{M}} I$. For M an even integer, denote the eigenvalues of W by $\{e^{jw_m}\}_{m=1}^{M}$. Let the first $M/2$ eigenvalues ($\{e^{jw_m}\}_{m=1}^{M/2}$) be chosen uniformly at random on the complex unit circle (i.e., we chose $\{w_m\}_{m=1}^{M/2}$ uniformly at random from $[0, 2\pi)$) and the other $M/2$ eigenvalues as the complex conjugates of these values (i.e., for $M/2 < m \leq M$, $e^{jw_m} = e^{-jw_{m-M/2}}$). Under these conditions, for a given RIP conditioning $\delta < 1$ and failure probability ρ, if*

$$M \geq C \frac{S}{\delta^2} \mu^2(\Psi) \log^4(N) \log(\rho^{-1}),$$

for a universal constant C, then for any s that is S-sparse (i.e., has no more than S non-zero entries)

$$(1-\delta) \leq \|\Phi\Psi s\|_2^2 / \|s\|_2^2 \leq (1+\delta)$$

with probability exceeding $1-\rho$.

We note that [111] shows that for the canonical basis ($\Psi = I$), the bounds for M can be tightened to $M \geq \max\left\{C \frac{S}{\delta^2} \log^4 N,\ C' \frac{S}{\delta^2} \log \rho^{-1}\right\}$ using a more complex

proof technique than we will employ here. For $\rho = \frac{1}{N}$, the result in [111] represents an improvement of a $\log(N)$ factor when restricted to only the canonical basis for Ψ.

While the proof of Theorem F.1.1 is fairly technical, the procedure follows very closely the proof of Theorem 8.1 from [111] on subsampled DTFT matrices. While the basic approach is the same, the novelty in our presentation is the incorporation of the sparsity basis Ψ and considerations for a real-valued connectivity matrix W.

Before beginning the proof of this theorem, we note that because U is assumed unitary, $\|\Phi\Psi s\|_2 = \|\widetilde{Z}F\Psi s\|_2$ for any signal s. Thus, it suffices to establish the conditioning properties of the matrix $\widehat{\Phi} := \widetilde{Z}F\Psi$. For the upcoming proof, it will be useful to write this matrix as a sum of rank-1 operators. The specific rank-1 operator that will be useful for our purposes is $X_l X_l^H$ with $X_l^H := F_l^H \Psi$, the l-th row of $F\Psi$, where $F_l^H := \left[1,\ e^{jw_l},\ \cdots,\ e^{jw_l(N-1)}\right] \in \mathbb{C}^N$ is the l-th row of F. Because of the way the "frequencies" $\{w_m\}$ are chosen, for any $l > \frac{M}{2}$, $X_l = X^*_{l-\frac{M}{2}}$. The l-th row of $\widehat{\Phi}$ is $\widetilde{z}_l X_l^H$ where $\widetilde{z}_l$ is the l-th diagonal entry of the diagonal matrix $\widetilde{Z}$, meaning that we can use the sum of rank-1 operators to write the decomposition $\widehat{\Phi}^H\widehat{\Phi} = \sum_{l=1}^{M} |\widetilde{z}_l|^2 X_l X_l^H$. If we define the random variable $\mathbf{B} := \widehat{\Phi}^H\widehat{\Phi} - I$ and the norm $|||\mathbf{B}||| := \sup_{y \text{ is } S\text{-sparse}} \frac{s^H \mathbf{B} s}{s^H s}$, we can equivalently say that $\widehat{\Phi}$ has RIP conditioning δ if

$$|||\mathbf{B}||| := \left|\left|\left|\widehat{\Phi}^H\widehat{\Phi} - I\right|\right|\right| = \left|\left|\left|\sum_{l=1}^{M} |\widetilde{z}_l|^2 X_l X_l^H - I\right|\right|\right| \leq \delta.$$

To aid in the upcoming proof, we make a few preliminary observations and rewrite the quantities of interest in some useful ways. First, because of the correspondences between the summands in $\widehat{\Phi}^H\widehat{\Phi}$ (i.e. $X_l = X^*_{l-M/2}$), we can rewrite $\widehat{\Phi}^H\widehat{\Phi}$ as

$$\widehat{\Phi}^H\widehat{\Phi} \ = \ \sum_{l=1}^{M/2} |\widetilde{z}_l|^2 X_l X_l^H + \sum_{l=1}^{M/2} |\widetilde{z}_l|^2 \left(X_l X_l^H\right)^*,$$

making clear the fact that there are only $\frac{M}{2}$ independent w_m's. Under the assumption

of Theorem F.1.1, $\widetilde{z}_l = \frac{1}{\sqrt{M}}$ for $l = 1, \cdots, M$. Therefore,

$$\mathbb{E}\left\{\sum_{l=1}^{M/2} |\widetilde{z}_l|^2 X_l X_l^H\right\} = \sum_{l=1}^{M/2} |\widetilde{z}_l|^2 \mathbb{E}\left\{X_l X_l^H\right\} = \sum_{l=1}^{M/2} \frac{1}{M} \Psi^H \mathbb{E}\left\{F_l F_l^H\right\} \Psi = \frac{1}{2} I,$$

where it is straightforward to check that $\mathbb{E}\left\{F_l F_l^H\right\} = I$. By the same reasoning, we also have $\mathbb{E}\left\{\sum_{l=1}^{M/2} |\widetilde{z}_l|^2 \left(X_l X_l^H\right)^*\right\} = \frac{1}{2} I$. This implies that we can rewrite $\mathbf{B}$ as

$$\begin{aligned}\mathbf{B} &= \sum_{l=1}^{M} \left(|\widetilde{z}_l|^2 X_l X_l^H\right) - I = \left(\sum_{l=1}^{M/2} |\widetilde{z}_l|^2 X_l X_l^H - \frac{1}{2} I\right) + \left(\sum_{l=1}^{M/2} |\widetilde{z}_l|^2 \left(X_l X_l^H\right)^* - \frac{1}{2} I\right) \\ &=: \mathbf{B}_1 + \mathbf{B}_2.\end{aligned}$$

The main proof of the theorem has two main steps. First, we will establish a bound on the moments of the quantity of interest $|||\mathbf{B}|||$ using Lemma F.1.1 below.[1] Next we will use these moments to derive a tail bound on $|||\mathbf{B}|||$ using Lemma A.1.2 in Section A.1, which will lead directly to the RIP statement we seek.

Lemma F.1.1. *Suppose $M \geq S$ and suppose we have a sequence of (fixed) vectors $Y_l \in \mathbb{C}^N$ for $l = 1, \cdots, M$ such that $K := \max_{l=1,\cdots,M} \|Y_l\|_\infty < \infty$. Let $\{\xi_l\}$ be a Rademacher sequence, i.e., a sequence of i.i.d. ± 1 random variables. Then for $p = 1$ and for $p \in \mathbb{R}$ and $p \geq 2$,*

$$\begin{aligned}&\left(\mathbb{E}\left\{\left|\left|\left|\sum_{l=1}^{M} \xi_l Y_l Y_l^H\right|\right|\right|^p\right\}\right)^{1/p} \\ &\leq C' C^{1/p} K \sqrt{p} \sqrt{S} \log(100S) \sqrt{\log(4N)\log(10M)} \sqrt{\left|\left|\left|\sum_{l=1}^{M} Y_l Y_l^H\right|\right|\right|},\end{aligned}$$

where C, C' are universal constants.

Armed with this notation and these lemmas, we now prove Theorem F.1.1:

Proof. We seek to show that under the conditions on M in Theorem F.1.1, $\mathbb{P}\{|||\mathbf{B}||| > \delta\} \leq \rho$. Since $\mathbf{B} = \mathbf{B}_1 + \mathbf{B}_2$ and $\{|||\mathbf{B}_1||| \leq \delta/2\} \cap \{|||\mathbf{B}_2||| \leq \delta/2\} \subset \{|||\mathbf{B}||| \leq \delta\}$, then,

$$\mathbb{P}\{|||\mathbf{B}||| > \delta\} \leq \mathbb{P}\{|||\mathbf{B}_1||| > \delta/2\} + \mathbb{P}\{|||\mathbf{B}_2||| > \delta/2\}.$$

[1] This is a reproduction of Lemma 2.4.1 in Section 2.4.3.

Thus, it will suffice to bound $\mathbb{P}\{|||\mathbf{B}_1||| > \delta/2\} \leq \rho/2$ since $\mathbf{B}_2 = \mathbf{B}_1^*$ implies that $\mathbb{P}\{|||\mathbf{B}_2||| > \delta/2\} \leq \rho/2$. In this presentation we let C, C' be some universal constant that may not be the same from line to line.

To begin, we use Lemma F.1.1 to bound $E_p := \left(\mathbb{E}\left\{|||\mathbf{B}_1|||^p\right\}\right)^{1/p}$ by setting $Y_l = \widetilde{z}_l^* X_l$ for $l = 1, \cdots, \frac{M}{2}$. To meet the conditions of Lemma F.1.1 we use a standard symmetrization manipulation (see Lemma A.1.1 of Section A.1). Specifically, we can write:

$$\begin{aligned} E_p = \left(\mathbb{E}\left\{|||\mathbf{B}_1|||^p\right\}\right)^{1/p} &\leq 2\left(\mathbb{E}\left\{\left|\left|\left|\sum_{l=1}^{M/2} \xi_l Y_l Y_l^H\right|\right|\right|^p\right\}\right)^{1/p} \\ &= 2\left(\mathbb{E}\left\{\left|\left|\left|\sum_{l=1}^{M/2} \xi_l |\widetilde{z}_l|^2 X_l X_l^H\right|\right|\right|^p\right\}\right)^{1/p}, \end{aligned}$$

where now the expectation is over the old random sequence $\{w_l\}$, together with a newly added Rademacher sequence $\{\xi_l\}$. Applying the law of iterated expectation and Lemma F.1.1, we have for $p \geq 2$:

$$\begin{aligned} E_p^p &:= \mathbb{E}\left\{|||\mathbf{B}_1|||^p\right\} \qquad (99) \\ &\leq 2^p \mathbb{E}\left\{\mathbb{E}\left\{\left|\left|\left|\sum_{l=1}^{M/2} \xi_l |\widetilde{z}_l|^2 X_l X_l^H\right|\right|\right|^p \mid \{w_l\}\right\}\right\} \\ &\leq \left(2C'C^{1/p}\sqrt{p}K\sqrt{S}\log(100S)\sqrt{\log(4N)\log(5M)}\right)^p \mathbb{E}\left\{\left|\left|\left|\sum_{l=1}^{M/2} |\widetilde{z}_l|^2 X_l X_l^H\right|\right|\right|^{p/2}\right\} \\ &\leq \left(C^{1/p}\sqrt{p}K\sqrt{C'S\log^4(N)}\right)^p \mathbb{E}\left\{\left(\left|\left|\left|\sum_{l=1}^{M/2}\left(|\widetilde{z}_l|^2 X_l X_l^H - \frac{1}{2}I\right)\right|\right|\right| + \frac{1}{2}|||I|||\right)^{p/2}\right\} \\ &\leq \left(C^{1/p}\sqrt{p}K\sqrt{C'S\log^4(N)}\right)^p \sqrt{\mathbb{E}\left\{\left(\left|\left|\left|\sum_{l=1}^{M/2}\left(|\widetilde{z}_l|^2 X_l X_l^H - \frac{1}{2}I\right)\right|\right|\right| + \frac{1}{2}\right)^p\right\}} \\ &\leq \left(C^{1/p}\sqrt{p}K\sqrt{C'S\log^4(N)}\right)^p \sqrt{\left(E_p + \frac{1}{2}\right)^p}. \end{aligned}$$

In the first line above, the inner expectation is over the Rademacher sequence $\{\xi_l\}$ (where we apply Lemma F.1.1) while the outer expectation is over the $\{w_l\}$. The

third line uses the triangle inequality for the $|||\cdot|||$ norm, the fourth line uses Jensen's inequality (Lemma A.1.4), and the fifth line uses triangle inequality for moments norm (i.e., $(\mathbb{E}\{|X+Y|^p\})^{1/p} \leq (\mathbb{E}\{|X|^p\})^{1/p} + (\mathbb{E}\{|Y|^p\})^{1/p}$). To get to $\log^4 N$ in the third line, we used our assumption that $N \geq M$, $N \geq S$ and $N \geq O(1)$ in Theorem F.1.1. Now using the definition of K from Lemma F.1.1, we can bound this quantity as:

$$K := \max_l \|Y_l\|_\infty = \max_l |\widetilde{z}_l| \|X_l\|_\infty = \frac{1}{\sqrt{M}} \max_l \|X_l\|_\infty = \frac{1}{\sqrt{M}} \max_{l,n} |\langle F_l, \Psi_n \rangle| \leq \frac{\mu(\Psi)}{\sqrt{M}}.$$

Therefore, we have the following implicit bound on the moments of the random variable of interest

$$E_p \quad \leq \quad C^{1/p}\sqrt{p}\sqrt{\frac{C'S\mu(\Psi)^2\log^4(N)}{M}}\sqrt{E_p + \frac{1}{2}}.$$

The above can be written as $E_p \leq a_p\sqrt{E_p + \frac{1}{2}}$, where $a_p = C^{1/p}\sqrt{p}\sqrt{\frac{4C'S\mu(\Psi)^2\log^4(N)}{M}}$. By squaring, rearranging the terms and completing the square, we have $E_p \leq \frac{a_p^2}{2} + a_p\sqrt{\frac{1}{2} + \frac{a_p^2}{4}}$. By assuming $a_p \leq \frac{1}{2}$, this bound can be simplified to $E_p \leq a_p$. Now, this assumption is equivalent to having an upper bound on the range of values of p:

$$\begin{aligned} a_p \leq \frac{1}{2} \quad &\Leftrightarrow \quad \sqrt{p} \leq \frac{1}{2C^{1/p}}\sqrt{\frac{M}{4C'S\mu(\Psi)^2\log^4(N)}} \\ &\Leftrightarrow \quad p \leq \frac{M}{16C^{2/p}C'S\mu(\Psi)^2\log^4(N)}. \end{aligned}$$

Hence, by using Lemma A.1.2 with $\alpha = \sqrt{\frac{C'S\mu(\Psi)^2\log^4(N)}{M}}$, $\beta = C$, $\gamma = 2$, $p_0 = 2$, and $p_1 = \frac{M}{16C^{2/p}C'S\mu(\Psi)^2\log^4(N)}$ we obtain the following tail bound for $u \in [\sqrt{2},\ \sqrt{p_1}]$:

$$\mathbb{P}\left\{ |||\mathbf{B}_1||| \geq e^{1/2}\sqrt{\frac{C'S\mu(\Psi)^2\log^4(N)}{M}}u \right\} \leq Ce^{-u^2/2}.$$

If we pick $\delta < 1$ such that

$$e^{1/2}\sqrt{\frac{C'S\mu(\Psi)^2\log^4(N)}{M}}u \leq \frac{\delta}{2} \tag{100}$$

and u such that

$$Ce^{-u^2/2} \leq \frac{\rho}{2} \quad \Leftrightarrow \quad u \geq \sqrt{2\log(2C\rho^{-1})},$$

then we have our required tail bound of $\mathbb{P}\{|||\mathbf{B}_1||| > \delta\} \leq \rho/2$. First, observe that (100) is equivalent to having

$$M \geq \frac{CS\mu(\Psi)^2 \log^4(N) \log(\rho^{-1})}{\delta^2}.$$

Also, because of the limited range of values u can take (i.e., $u \in [\sqrt{2}, \sqrt{p_1}]$), we require that

$$\begin{aligned} \sqrt{2\log(2C\rho^{-1})} &\leq \sqrt{\frac{M}{16C^{2/p}C'S\mu(\Psi)^2\log^4(N)}} = \sqrt{p_1} \\ \Leftrightarrow \quad M &\geq CS\mu(\Psi)^2\log^4(N)\log(\rho^{-1}), \end{aligned}$$

which, together with the earlier condition on M, completes the proof. □

F.2 Bounded Error due to Stimulus History

In this section, we derive the bound (47) in Section 6.4. The approach we take is to bound the individual components of (8) pertaining to the ℓ_1 recovery errors. Specifically, we will bound the noise term (the first term in (8)) by the component of the input history that is beyond the attempted recovery, and we will bound the signal approximation term (the second term in (8)) by the quality of the signal recovery possible in the attempted recovery length. In this way we can observe how different properties of the system and input sequence affect signal recovery.

To bound the first term in (8) (i.e., the omission errors due to inputs beyond the recovery window), we first write the current state at any time N^* as

$$x[N^*] = \sum_{n=0}^{N^*} W^{N^*-n} z s[n].$$

We only wish to recover the past $N \leq N^*$ time steps, so we break up the summation into components of the current state due to "signal" (i.e., signal we attempt to recover)

and "noise" (i..e, older signal we omit from the recovery):

$$\begin{aligned} x[N^*] &= \sum_{n=N^*-N+1}^{N^*} W^{N^*-n} zs[n] + \sum_{n=0}^{N^*-N} W^{N^*-n} zs[n] \\ &= \sum_{n=N^*-N+1}^{N^*} W^{N^*-n} zs[n] + \epsilon \\ &= \Phi s + \epsilon. \end{aligned}$$

From here we can see that the first summation is the matrix multiply Φs as is discussed in the paper. The second summation here, ϵ, takes the place of a noise term in the recovery. We can further analyze the effect of this noise term by understanding that ϵ is a bounded noise term for well behaved input sequences $s[n]$ (in fact, all that is needed is that the maximum value or the expected value and variance are reasonably bounded) when the eigenvalues of W are of magnitude $q \leq 1$. We can explicitly calculate the worst case scenario bounds on the norm of ϵ,

$$\begin{aligned} \left\| \sum_{n=0}^{N^*-N} W^{N^*-n} zs[n] \right\|_2 &\leq \left\| \sum_{n=0}^{N^*-N} U \left(qD\right)^{N^*-n} U^{-1} zs[n] \right\|_2 \\ &\leq \left\|U\right\|_2 \left\| \sum_{n=0}^{N^*-N} \left(qD\right)^{N^*-n} U^{-1} zs[n] \right\|_2, \end{aligned}$$

where $D = \mathrm{diag}(d_1, \ldots, d_M)$ is the diagonal matrix containing the normalized eigenvalues of W. If we assume that z is chosen as mentioned as in Section 6.3 so that $U^{-1}z = \left(1/\sqrt{M}\right)\mathbf{1}$, the eigenvalues of W are uniformly spread around a complex

circle of radius q, and that $s[n] \leq s_{\max}$ for all n, then we can bound this quantity as

$$\begin{aligned}
\left\| \sum_{n=0}^{N^*-N} W^{N^*-n} z s[n] \right\|_2 &\leq \frac{s_{\max}}{\sqrt{M}} \|U\|_2 \left\| \sum_{n=0}^{N^*-N} (qD)^{N^*-n} \mathbf{1} \right\|_2 \\
&= \frac{s_{\max}}{\sqrt{M}} \|U\|_2 \left\| \begin{bmatrix} \sum_{n=0}^{N^*-N} q^{N^*-n} d_1^{N^*-n} \\ \vdots \\ \sum_{n=0}^{N^*-N} q^{N^*-n} d_M^{N^*-n} \end{bmatrix} \right\|_2 \\
&= \frac{s_{\max}}{\sqrt{M}} \|U\|_2 \sqrt{\sum_{k=1}^{M} \left| \sum_{n=0}^{N^*-N} q^{N^*-n} d_k^{N^*-n} \right|^2} \\
&\leq s_{\max} \|U\|_2 \left| \sum_{n=0}^{N^*-N} q^{N^*-n} \right| \\
&\leq s_{\max} \|U\|_2 \left| \frac{q^N - q^{N^*}}{1-q} \right|
\end{aligned}$$

where d_k is the k^{th} normalized eigenvalue of W. In the limit of large input signal lengths ($N^* \to \infty$), we have $N^* \gg N$ and so $q^N \gg q^{N^*}$, which leaves the approximate expression

$$\|\epsilon\|_2 \leq s_{\max} \|U\|_2 \left| \frac{q^N}{1-q} \right|.$$

To bound the second term in (8) (i.e., the signal approximation errors due to imperfect recovery), we must characterize the possible error between the signal (which is S-sparse) and the approximation to the signal with the S^* largest coefficients. In the worst case scenario, there are $S - S^* + 1$ coefficients that cannot be guaranteed to be recovered by the RIP conditions, and these coefficients all take the maximum value $s_{\max}$. In this case, we can bound the signal approximation error:

$$\begin{aligned}
\frac{\beta}{\sqrt{S^*}} \|s - s_{S^*}\|_1 &\leq \frac{\beta}{\sqrt{S^*}} \sum_{n=S^*+1}^{S} |q^n s_{\max}| \\
&= \frac{\beta s_{\max}}{\sqrt{S^*}} \left(\frac{q^{S^*} - q^S}{1-q} \right).
\end{aligned}$$

F.3 Gaussian Uncertainties in Feed-forward Vector

In this section, we extend the RIP analysis of Appendix F.1 to the case when z is chosen to be a Gaussian i.i.d. vector. The RIP analysis is summarized in the following theorem:

Theorem F.3.1. *Suppose $N \geq M$, $N \geq S$ and $N \geq O(1)$. Let the entries of z be i.i.d. zero-mean Gaussian random variables with variance $\frac{1}{M}$, and the eigenvector matrix U be unitary. For M an even integer, denote the eigenvalues of W by $\{e^{jw_m}\}_{m=1}^{M}$. Let the first $M/2$ eigenvalues ($\{e^{jw_m}\}_{m=1}^{M/2}$) be chosen uniformly at random on the complex unit circle (i.e., we chose $\{w_m\}_{m=1}^{M/2}$ uniformly at random from $[0, 2\pi)$) and the other $M/2$ eigenvalues as the complex conjugates of these values. Then, for a given RIP conditioning δ and failure probability $N^{-\log^4 N} \leq \rho \leq \frac{1}{e}$, if*

$$M \geq C\frac{S}{\delta^2}\mu^2(\Psi)\log^5(N)\log(\rho^{-1}),$$

Φ satisfies RIP-(S, δ) with probability exceeding $1 - \rho$ for a universal constant C.

It is unfortunate that with the additional randomness in the feed-forward vector, the same proof procedure as in Theorem F.1.1 cannot be used. In the proof of Theorem F.1.1, we showed that the random variable $|||Z_1|||$ has p-th moments that scale like $\alpha\beta^{1/p}p^{1/2}$ (through Lemma F.1.1) for a range of p which suggests that it has a subgaussian tail (i.e., $\mathbb{P}\{|||Z_1||| > u\} \leq Ce^{-u^2/2}$) for a range of deviations u. We then used this tail bound to bound the probability that $|||Z_1|||$ exceeds a fixed conditioning δ. With Gaussian uncertainties in the feed-forward vector z, Lemma F.1.1 will not yield the required subgaussian tail but instead gives us moments estimates that result in sub-optimal scaling of M with respect to N. Therefore, we will instead follow the proof procedure of Theorem 16 from [133] that will yield the better measurement rate given in Theorem F.3.1 above.

Let us begin by recalling a few notations from the proof of Theorem F.1.1 and by introducing further notations that will simplify our exposition later. First, recall

that we let X_l^H be the l-th row of $F\Psi$. Thus, the l-th row of our matrix of interest $\widehat{\Phi} = \widetilde{Z}F\Psi$ is $\widetilde{z}_l X_l^H$ where $\widetilde{z}_l$ is the l-th diagonal entry of the diagonal matrix $\widetilde{Z}$. Whereas before, $\widetilde{z}_l = \frac{1}{\sqrt{M}}$ for any $l = 1, \cdots, M$, here it will a random variable. To understand the resulting distribution of $\widetilde{z}_l$, first note that for the connectivity matrix W to be real, we need to assume that the second $\frac{M}{2}$ columns of U are complex conjugates of the first $\frac{M}{2}$ columns. Thus, we can write $U = [U_R \mid U_R] + j\,[U_I \mid -U_I]$, where $U_R, U_I \in \mathbb{R}^{M \times \frac{M}{2}}$. Because $U^H U = I$, we can deduce that $U_R^T U_I = I$ and that the ℓ_2 norms of the columns of both U_R and U_I are $\frac{1}{\sqrt{2}}$.[2] With these matrices U_R, U_I, let us re-write the random vector $\widetilde{z}$ to illustrate its structure. Consider the matrix $\widehat{U} := [U_R \mid U_I] \in \mathbb{R}^{M \times M}$, which is a scaled unitary matrix (because we can check that $\widehat{U}^T\widehat{U} = \frac{1}{2}I$). Next, consider the random vector $\widehat{z} := \widehat{U}^T z$. Because $\widehat{U}$ is (scaled) unitary and z is composed of i.i.d. zero-mean Gaussian random variables of variance $\frac{1}{M}$, the entries of $\widehat{z}$ are also i.i.d. zero-mean Gaussian random variables, but now with variance $\frac{1}{2M}$. Then, from our definition of U in terms of U_R and U_I, for any $l \leq \frac{M}{2}$, we have $\widetilde{z}_l = \widehat{z}_l + j\widehat{z}_{l+\frac{M}{2}}$ and for $l > \frac{M}{2}$, we have $\widetilde{z}_l = \widehat{z}_{l-\frac{M}{2}} - j\widehat{z}_l$. This clearly shows that each of the *first* $\frac{M}{2}$ entries of $\widetilde{z}$ is made up of 2 i.i.d. random variables (one being the real component, the other imaginary), and that the other $\frac{M}{2}$ entries are just complex conjugates of the first $\frac{M}{2}$. Because of this, for $l \leq \frac{M}{2}$, $|\widetilde{z}_l|^2 = |\widetilde{z}_{l+\frac{M}{2}}|^2 = \widehat{z}_l^2 + \widehat{z}_{l+\frac{M}{2}}^2$ is the sum of squares of 2 i.i.d. Gaussian random variables.

From the proof of Theorem F.1.1, we also denoted

$$Z \;:=\; \widehat{\Phi}^H\widehat{\Phi} - I \;=\; \left(\sum_{l=1}^{M/2} |\widetilde{z}_l|^2 X_l X_l^H - \frac{1}{2}I\right) + \left(\sum_{l=1}^{M/2} |\widetilde{z}_l|^2 \left(X_l X_l^H\right)^* - \frac{1}{2}I\right) \;=:\; Z_1 + Z_2.$$

[2] This can be shown by writing

$$\begin{aligned} U^H U &= \left(\begin{bmatrix} U_R^T \\ U_R^T \end{bmatrix} - j\begin{bmatrix} U_I^T \\ -U_I^T \end{bmatrix}\right)([U_R \mid U_R] + j\,[U_I \mid -U_I]) \\ &= \left(\begin{bmatrix} U_R^T U_R & U_R^T U_R \\ U_R^T U_R & U_R^T U_R \end{bmatrix} + \begin{bmatrix} U_I^T U_I & -U_I^T U_I \\ -U_I^T U_I & U_I^T U_I \end{bmatrix}\right) + j\left(\begin{bmatrix} U_R^T U_I & -U_R^T U_I \\ U_R^T U_I & -U_R^T U_I \end{bmatrix} + \begin{bmatrix} U_I^T U_R & U_I^T U_R \\ -U_I^T U_R & -U_I^T U_R \end{bmatrix}\right). \end{aligned}$$

Then by equating the above to $I + j0$, we arrive at our conclusion.

It is again easy to check that $\mathbb{E}\left\{\sum_{l=1}^{M/2}\left(|\widetilde{z}_l|^2 X_l X_l^H\right)\right\} = \mathbb{E}\left\{\sum_{l=1}^{M/2}\left(|\widetilde{z}_l|^2\left(X_l X_l^H\right)^*\right)\right\} = \frac{1}{2}I$. Finally, $\widehat{\Phi}$ has RIP conditioning δ whenever $|||Z||| \leq \delta$ with $|||Z||| := \sup_{y \text{ is } S\text{-sparse}} \frac{y^H Z y}{y^H y}$.

Before moving on to the proof, we first present a lemma regarding the random sequence $|z_l|^2$ that will be useful in the sequel.

Lemma F.3.1. *Suppose for $l = 1, \cdots, \frac{M}{2}$, $|\widetilde{z}_l|^2 = \widehat{z}_l^2 + \widehat{z}_{l+M/2}^2$ where $\widehat{z}_l$ for $l = 1, \cdots, M$ is a sequence of i.i.d. zero-mean Gaussian random variables of variance $\frac{1}{2M}$. Also suppose that $\rho \leq 1$ is a fixed probability. For the random variable $\max_{l=1,\cdots,M/2}|\widetilde{z}_l|^2$, we have the following bounds on the expected value and tail probability of this extreme value:*

$$\mathbb{E}\left\{\max_{l=1,\cdots,M/2}|\widetilde{z}_l|^2\right\} \leq \frac{1}{M}\left(\log\left(\frac{C_1 M}{2}\right)+1\right), \quad (101)$$

$$\mathbb{P}\left\{\max_{l=1,\cdots,M/2}|\widetilde{z}_l|^2 > \frac{C_2\log\left(C_2' M\rho^{-1}\right)}{M}\right\} \leq \rho. \quad (102)$$

Proof. To ease notation, every index l used as a variable for a maximization will be taken over the set $l = 1, \ldots, \frac{M}{2}$ without explicitly writing the index set. To calculate $\mathbb{E}\{\max_l |\widetilde{z}_l|^2\}$, we use the following result that allows us to bound the expected value of a positive random variable by its tail probability (see Lemma A.1.3 of Section A.1):

$$\mathbb{E}\left\{\max_l |\widetilde{z}_l|^2\right\} = \int_0^\infty \mathbb{P}\left\{\max_l |\widetilde{z}_l|^2 > u\right\} du. \quad (103)$$

Using the union bound, we have the estimate $\mathbb{P}\{\max_l |\widetilde{z}_l|^2 > u\} \leq \frac{M}{2}\mathbb{P}\{|\widetilde{z}_1|^2 > u\}$ (since the $|\widetilde{z}_l|^2$ are identically distributed). Now, because $|\widetilde{z}_1|^2$ is a sum of squares of two Gaussian random variables and thus is a (generalized) χ^2 random variable with 2 degrees of freedom (which we shall denote by χ_2),[3] we have

$$\mathbb{P}\left\{|\widetilde{z}_1|^2 > u\right\} = \mathbb{P}\{\chi_2 > 2Mu\} = \frac{1}{2\Gamma(1)} e^{\frac{-2Mu}{2}} = C_1 e^{-Mu},$$

[3]The pdf of a χ^2 random variable χ_q with q degrees of freedom is given by $p(x) = \frac{1}{2^{q/2}\Gamma(q/2)} x^{q/2-1} e^{-x/2}$. Therefore, it's tail probability can be obtained by integration: $\mathbb{P}\{\chi_q > u\} = \int_u^\infty p(x)dx$.

where $\Gamma(\cdot)$ is the Gamma function and the $2Mu$ appears instead of u in the exponential is because of the standardization of the Gaussian random variables (initially of variance $\frac{1}{2M}$). To proceed, we break the integral in (103) into 2 parts. To do so, notice that if $u < \frac{1}{M}\log\left(\frac{C_1 M}{2}\right)$, then the trivial upper bound of $\mathbb{P}\{\max_l |\widetilde{z}_l|^2 > u\} \leq 1$ is a better estimate than $\frac{C_1 M}{2} e^{-Mu}$. In other words, our estimate for the tail bound of $\max_l |\widetilde{z}_l|^2$ is not very good for small u but gets better with increasing u. Therefore, we have

$$\begin{aligned}
\mathbb{E}\left\{\max_l |\widetilde{z}_l|^2\right\} &\leq \int_0^{\frac{1}{M}\log\left(\frac{C_1 M}{2}\right)} 1 \, du + \int_{\frac{1}{M}\log\left(\frac{C_1 M}{2}\right)}^{\infty} \frac{C_1 M}{2} e^{-Mu} \, du \\
&= \frac{1}{M}\log\left(\frac{C_1 M}{2}\right) - \frac{C_1 M}{2}\left[\frac{1}{M}e^{-Mu}\right]\Bigg|_{\frac{1}{M}\log\left(\frac{C_1 M}{2}\right)}^{\infty} \\
&= \frac{1}{M}\log\left(\frac{C_1 M}{2}\right) + \frac{C_1}{2}e^{-\log\left(\frac{C_1 M}{2}\right)} = \frac{1}{M}\left(\log\left(\frac{C_1 M}{2}\right) + 1\right).
\end{aligned}$$

This is the bound in expectation that we seek for in (103).

In the second part of the proof that follows, C, C' denote universal constants. Essentially, we will want to apply Lemma A.1.2 that is used in Appendix F.1 to obtain our tail bound. In the lemma, the tail bound of a random variable X can be estimated once we know the moments of X. Therefore, we require the moments of the random variable $\max_l |\widetilde{z}_l|^2$. For this, for any $p > 0$, we use the following simple estimate:

$$\mathbb{E}\left\{\max_l |\widetilde{z}_l|^{2p}\right\} \leq \frac{M}{2}\max_l \mathbb{E}\left\{|\widetilde{z}_l|^{2p}\right\} = \frac{M}{2}\mathbb{E}\left\{|\widetilde{z}_1|^{2p}\right\}, \tag{104}$$

where the first step comes from writing the expectation as an integral of the cumulative distribution (as seen in (103)) and taking the union bound, and the second step comes from the fact that the $|\widetilde{z}_l|^2$ are identically distributed. Now, $|\widetilde{z}_1|^2$ is a sub-exponential random variable since it is a sum of squares of Gaussians [136].[4]

[4] A sub-exponential random variable is a random variable whose tail probability scale like $\exp^{-Cu}$ for some constant C. Thus, a χ^2 random variable is a specific instance of a sub-exponential random variable.

Therefore, for any $p > 0$, it's p-th moment can be bounded by

$$\mathbb{E}\left\{|\widetilde{z}_1|^{2p}\right\}^{1/p} \leq \frac{C'}{M} C^{1/p} p,$$

where the division by M comes again from the variance of the Gaussian random variables that make up $|\widetilde{z}_1|^2$. Putting this bound with (104), we have the following estimate for the p-th moments of $\max_l |\widetilde{z}_l|^2$:[5]

$$\mathbb{E}\left\{\max_l |\widetilde{z}_l|^{2p}\right\}^{1/p} \leq \frac{C'}{M}\left(\frac{CM}{2}\right)^{1/p} p.$$

Therefore, by Lemma A.1.2 with $\alpha = \frac{C'}{M}$, $\beta = \frac{CM}{2}$, and $\gamma = 1$, we have

$$\mathbb{P}\left\{\max_l |\widetilde{z}_l|^2 > \frac{eC'u}{M}\right\} \leq \frac{CM}{2} e^{-u}.$$

By choosing $u = \log\left(\frac{CM}{2}\rho^{-1}\right)$, we have our desired tail bound. □

Armed with this lemma, we can now turn our attention to the main proof. As stated earlier, this follows essentially the same form as [133] with the primary difference of including the results from Lemma F.3.1. As before, because $\mathbb{P}\{\|\!|Z|\!\| > \delta\} \leq \mathbb{P}\{\|\!|Z_1|\!\| > \delta/2\} + \mathbb{P}\{\|\!|Z_2|\!\| > \delta/2\}$ with $Z_2 = Z_1^*$, we just have to consider bounding the tail bound $\mathbb{P}\{\|\!|Z_1|\!\| > \delta/2\}$. This proof differs from that in Appendix F.1 in that here, we will first show that $\mathbb{E}\{\|\!|Z_1|\!\|\}$ is small when M is large enough and then show that Z_1 does not differ much from $\mathbb{E}\{\|\!|Z_1|\!\|\}$ with high probability.

F.3.1 Expectation

In this section, we will show that $\mathbb{E}\{\|\!|Z_1|\!\|\}$ is small. This will basically follow from Lemma F.1.1 in Appendix F.1 and (101) in Lemma F.3.1. To be precise, the remainder of this section is to prove:

Theorem F.3.2. *Choose any $\delta' \leq \frac{1}{2}$. If $M \geq \frac{C_3 S\mu(\Psi)^2 \log^5 N}{\delta'^2}$, then $\mathbb{E}\{\|\!|Z|\!\|\} \leq \delta'$.*

[5] We remark that this bound gives a worse estimate for the expected value as that calculated before because of the crude bound given by (104).

Proof. Again, C is some universal constant that may not be the same from line to line. We follow the same symmetrization step found in the proof in Appendix F.1 to arrive at:

$$E := \mathbb{E}\left\{|||Z_1|||\right\} \;\leq\; 2\mathbb{E}\left\{\mathbb{E}\left\{\left|\left|\left|\sum_{l=1}^{M/2}\xi_l|\widetilde{z}_l|^2 X_l X_l^H\right|\right|\right| |\{w_l\},\widetilde{z}\right\}\right\},$$

where the outer expectation is over the Rademacher sequence $\{\xi_l\}$ and the inner expectation is over the random "frequencies" $\{w_l\}$ and feed-forward vector $\widetilde{z}$. As before, for $l = 1, \cdots, \frac{M}{2}$, we set $Y_l = \widetilde{z}_l^* X_l$. Observe that by definition $K := \max_{l=1,\cdots,M/2}\|Y_l\|_\infty = \max_l|\widetilde{z}_l|\|X_l\|_\infty$ and thus is a random variable. We then use Lemma F.1.1 with $p = 1$ to get

$$\begin{aligned}
E \;&\leq\; 2C\sqrt{S}\log(100S)\sqrt{\log(4N)\log(5M)}\mathbb{E}\left\{K\sqrt{\left|\left|\left|\sum_{l=1}^{M/2}|\widetilde{z}_l|^2 X_l X_l^H\right|\right|\right|}\right\} \\
&\leq\; \sqrt{4CS\log^4(N)}\sqrt{\mathbb{E}\left\{K^2\right\}}\sqrt{\mathbb{E}\left\{\left|\left|\left|\sum_{l=1}^{M/2}|\widetilde{z}_l|^2 X_l X_l^H\right|\right|\right|\right\}} \\
&\leq\; \sqrt{4CS\log^4(N)}\sqrt{\mathbb{E}\left\{K^2\right\}}\sqrt{E+\frac{1}{2}},
\end{aligned}$$

where the second line uses the Cauchy-Schwarz inequality for expectations and the third line uses triangle inequality. Again, to get to $\log^4 N$ in the second line, we used our assumption that $N \geq M$, $N \geq S$ and $N \geq O(1)$ in Theorem F.3.1. It therefore remains to calculate $\mathbb{E}\{K^2\}$. Now, $K = \max_l|\widetilde{z}_l|\|X_l\|_\infty \leq \max_l|\widetilde{z}_l|\max_l\|X_l\|_\infty$. First, we have $\max_l\|X_l\|_\infty = \max_{l,n}|\langle F_l, \Psi_n\rangle| \leq \mu(\Psi)$. Next, (101) in Lemma F.3.1 tells us that $\mathbb{E}\left\{\max_{l=1,\cdots,M/2}|\widetilde{z}_l|^2\right\} \leq \frac{1}{M}\left(\log\left(\frac{C_1 M}{2}\right)+1\right)$. Thus, we have $\mathbb{E}\{K^2\} \leq \frac{\mu(\Psi)^2}{M}\left(\log\left(\frac{C_1 M}{2}\right)+1\right)$. Putting everything together, we have

$$E = \mathbb{E}\left\{|||Z_1|||\right\} \leq \sqrt{\frac{CS\log^4(N)\left(\log\left(\frac{C_1 M}{2}\right)+1\right)\mu(\Psi)^2}{M}}\sqrt{E+\frac{1}{2}}.$$

Now, the above can be written as $E \leq a\sqrt{E+\frac{1}{2}}$, where $a = \sqrt{\frac{CS\log^4(N)\left(\log\left(\frac{C_1 M}{2}\right)+1\right)\mu(\Psi)^2}{M}}$. By squaring it, rearranging the terms and completing the squares, we have $E \leq$

$\frac{a^2}{2} + a\sqrt{\frac{1}{2} + \frac{a^2}{4}}$. By supposing $a \leq \frac{1}{2}$, this can be simplified as $E \leq a$. To conclude, let us choose M such that $a \leq \delta'$ where $\delta' \leq \frac{1}{2}$ is our pre-determined conditioning (which incidentally fulfills our previous assumption that $a \leq \frac{1}{2}$). By applying the formula for a, we have that if $M \geq \frac{C_3 S\mu(\Psi)^2 \log^5(N)}{\delta'^2}$, then $E \leq \delta'$. □

F.3.2 Tail Probability

To give a probability tail bound estimate to Z_1, we use the following lemma found in [111, 133]:

Lemma F.3.2. *Suppose Y_l for $l = 1, \cdots, M$ are independent, symmetric random variables such that $\|\|Y_l\|\| \leq K < \infty$ almost surely. Let $Y = \sum_{l=1}^{M} Y_l$. Then for any $u, t > 1$, we have*

$$\mathbb{P}\{\|\|Y\|\| > C(u\mathbb{E}\{\|\|Y\|\|\} + tK)\} \leq e^{-u^2} + e^{-t}.$$

The goal of this section is to prove:

Theorem F.3.3. *Pick any $\delta \leq \frac{1}{2}$ and suppose $N^{-\log^4(N)} \leq \rho \leq \frac{1}{e}$. Suppose $M \geq \frac{C_4 S\mu(\Psi)^2 \log^5 N \log \rho^{-1}}{\delta^2}$, then $\mathbb{P}\{\|\|Z_1\|\| > \delta\} \leq 8\rho$.*

Proof. To use Lemma F.3.2, we want Y_l to look like the summands of

$$Z_1 = \sum_{l=1}^{M/2} \left(|\widetilde{z}_l|^2 X_l X_l^H - \mathbb{E}\left\{ |\widetilde{z}_l|^2 X_l X_l^H \right\} \right).$$

However, this poses several problems. First, they are *not* symmetric and thus, we need to symmetrize it by defining

$$\begin{aligned} \widetilde{Y}_l &= |\widetilde{z}_l|^2 X_l X_l^H - |\widetilde{z}_l'|^2 X_l' (X_l')^H \\ &\sim \xi_l \left(|\widetilde{z}_l|^2 X_l X_l^H - |\widetilde{z}_l'|^2 X_l' (X_l')^H \right) \end{aligned}$$

where $\widetilde{z}', X_l'$ are independent copies of $\widetilde{z}$ and X_l respectively, and ξ_l is an independent Rademacher sequence. Here, the relation $X \sim Y$ for two random variables X, Y

means that X has the same distribution as Y. To form $\widetilde{Y}_l$, what we have done is take each summand of Z_1 and take it's difference with an independent copy of itself. Because $\widetilde{Y}_l$ is symmetric, adding a Rademacher sequence does not change its distribution and this sequence is only introduced to resolve a technicality that will arise later on. If we let $\widetilde{Y} := \sum_{l=1}^{M/2} \widetilde{Y}_l$, then the random variables $\widetilde{Y}$ (symmetrized) and Z_1 (un-symmetrized) are related via the following estimates [111]:

$$\mathbb{E}\left\{\left|\left|\left|\widetilde{Y}\right|\right|\right|\right\} \;\leq\; 2\mathbb{E}\left\{|||Z_1|||\right\}, \tag{105}$$

$$\mathbb{P}\{|||Z_1||| > 2\mathbb{E}\left\{|||Z_1|||\right\} + u\} \;\leq\; 2\mathbb{P}\left\{\left|\left|\left|\widetilde{Y}\right|\right|\right| > u\right\}. \tag{106}$$

However, a second condition imposed on Y_l in Lemma F.3.2 is that $|||Y_l||| \leq K < \infty$ almost surely. Because of the unbounded nature of the Gaussian random variables $\widetilde{z}_l$ and $\widetilde{z}'_l$ in $\widetilde{Y}_l$, this condition is not met. Therefore, we need to define a Y_l that is conditioned on the event that these Gaussian random variables are bounded. To do so, define the following event:

$$F = \left\{\max\left\{\max_l |\widetilde{z}_l|^2,\ \max_l |\widetilde{z}'_l|^2\right\} \leq \frac{C_2 \log\left(C_2' M\rho^{-1}\right)}{M}\right\}.$$

Using (102) in Lemma F.3.1, we can calculate $\mathbb{P}\{F^c\}$, where F^c is the complementary event of F:

$$\begin{aligned}
\mathbb{P}\{F^c\} &= \mathbb{P}\left\{\max\left\{\max_l |\widetilde{z}_l|^2,\ \max_l |\widetilde{z}'_l|^2\right\} > \frac{C_2 \log\left(C_2' M\rho^{-1}\right)}{M}\right\} \\
&\leq \mathbb{P}\left\{\max_l |\widetilde{z}_l|^2 > \frac{C_2 \log\left(C_2' M\rho^{-1}\right)}{M}\right\} + \mathbb{P}\left\{\max_l |\widetilde{z}'_l|^2 > \frac{C_2 \log\left(C_2' M\rho^{-1}\right)}{M}\right\} \leq 2\rho.
\end{aligned}$$

Conditioned on event F, the $|||\cdot|||$ norm of $\widetilde{Y}_l$ is well-bounded:

$$\begin{aligned}
\left|\left|\left|\widetilde{Y}_l\right|\right|\right| &= \left|\left|\left||\widetilde{z}_l|^2 X_l X_l^H - |\widetilde{z}'_l|^2 X'_l (X'_l)^H\right|\right|\right| \;\leq\; 2\max\left\{\max_l |\widetilde{z}_l|^2,\ \max_l |\widetilde{z}'_l|^2\right\} |||X_l X_l^H||| \\
&= \frac{2C_2 \log\left(C_2' M\rho^{-1}\right)}{M} \sup_{y \text{ is } S\text{-sparse}} \left\{\frac{y^H X_l X_l^H y}{y^H y}\right\} \\
&\leq \frac{2C_2 \log\left(C_2' M\rho^{-1}\right)}{M} \sup_{y \text{ is } S\text{-sparse}} \left\{\|X_l\|_\infty^2 \frac{\|y\|_1^2}{\|y\|_2^2}\right\} \\
&\leq \frac{2SC_2 \log\left(C_2' M\rho^{-1}\right)}{M} \max_l \|X_l\|_\infty^2 \;\leq\; \frac{CS\mu(\Psi)^2 \log\left(C_2' M\rho^{-1}\right)}{M} := K,
\end{aligned}$$

where in the last line we used the fact that the ratio between the ℓ_1 and ℓ_2 norms of an S-sparse vector is S, and the estimate we derived for $\max_l \|X_l\|_\infty^2$ in Appendix F.1.

We now define a new random variable that is a truncated version of $\widetilde{Y}_l$ which takes for value 0 whenever we fall under event F^c, i.e.,

$$Y_l := \widetilde{Y}_l \, \mathbb{I}_F = \xi_l \left(|\widetilde{z}_l|^2 X_l X_l^H - |\widetilde{z}_l'|^2 X_l' (X_l')^H \right) \mathbb{I}_F,$$

where $\mathbb{I}_F$ is the indicator function of event F_l. If we define $Y = \sum_{l=1}^{M/2} Y_l$, then the random variables Y (truncated) and $\widetilde{Y}$ (un-truncated) are related by [133] (see also Lemma 1.4.3 of [47])

$$\mathbb{P}\left\{ \left| \left| \left| \widetilde{Y} \right| \right| \right| > u \right\} \leq \mathbb{P}\{ |||Y||| > u \} + \mathbb{P}\{F^c\}. \tag{107}$$

When $\widetilde{z}, \widetilde{z}', X_l, X_l'$ are held constant so only the Rademacher sequence ξ_l is random, then the contraction principle [88, 133] tells us that $\mathbb{E}\{|||Y|||\} \leq \mathbb{E}\left\{ \left| \left| \left| \widetilde{Y} \right| \right| \right| \right\}$. Note that the sole reason for introducing the Rademacher sequences is for this use of the contraction principle. As this holds point-wise for all $\widetilde{z}, \widetilde{z}', X_l, X_l'$, we have

$$\mathbb{E}\{|||Y|||\} \leq \mathbb{E}\left\{ \left| \left| \left| \widetilde{Y} \right| \right| \right| \right\}. \tag{108}$$

We now have all the necessary ingredients to apply Lemma F.3.2. First, by choosing $\delta' \leq \frac{1}{2}$, from Theorem F.3.2, we have that $\mathbb{E}\{|||Z|||\} \leq \delta'$ whenever $M \geq \frac{C_3 S \mu(\Psi)^2 \log^5 N}{\delta'^2}$. Thus, by chaining (108) and (105), we have

$$\mathbb{E}\{|||Y|||\} \leq \mathbb{E}\left\{ \left| \left| \left| \widetilde{Y} \right| \right| \right| \right\} \leq 2\mathbb{E}\{|||Z_1|||\} \leq 2\delta'.$$

Also, with this choice of M, we have

$$K = \frac{C S \mu(\Psi)^2 \log(C_2' M \rho^{-1})}{M} \leq \frac{C \delta'^2 \log(C_2' M \rho^{-1})}{\log^5 N}.$$

Using these estimates for K and $\mathbb{E}\{|||Y|||\}$, and choosing $u = \sqrt{\log \rho^{-1}}$ and $t = \log \rho^{-1}$, Lemma F.3.2 says that

$$\mathbb{P}\left\{ |||Y||| > C' \left(2\delta' \sqrt{\log \rho^{-1}} + \frac{C \delta'^2 \log(C_2' M \rho^{-1}) \log \rho^{-1}}{\log^5 N} \right) \right\} \leq 2\rho.$$

Then, using the relation between the tail probabilities of Y and $\widetilde{Y}$ (107) together with our estimate for $\mathbb{P}\{F^c\}$, we have

$$\mathbb{P}\left\{ \left|\!\left|\!\left| \widetilde{Y} \right|\!\right|\!\right| > C' \left(2\delta' \sqrt{\log \rho^{-1}} + \frac{C\delta'^2 \log\left(C_2' M \rho^{-1}\right) \log \rho^{-1}}{\log^5 N} \right) \right\} \leq 2\rho + \mathbb{P}\{F^c\} \leq 4\rho.$$

Finally, using the relation between the tail probabilities of $\widetilde{Y}$ and Z (106), we have

$$\mathbb{P}\left\{ \left|\!\left|\!\left| Z_1 \right|\!\right|\!\right| > 2\delta' + 2C'\delta' \sqrt{\log \rho^{-1}} + \frac{CC'\delta'^2 \log\left(C_2' M \rho^{-1}\right) \log \rho^{-1}}{\log^5 N} \right\} \leq 8\rho,$$

where we used the fact that $\mathbb{E}\left\{ \left|\!\left|\!\left| Z_1 \right|\!\right|\!\right| \right\} \leq \delta'$. Then, for a pre-determined conditioning $\delta \leq \frac{1}{2}$, pick $\delta' = \frac{\delta}{3C''\sqrt{\log \rho^{-1}}}$ for a constant C'' which will be chosen appropriately later. With this choice of δ' and with our assumptions that $\delta \leq \frac{1}{2}$ and $\rho \leq \frac{1}{e}$, the three terms in the tail bound becomes

$$\begin{aligned}
2\delta' &= \frac{\delta}{3C''\sqrt{\log \rho^{-1}}} \leq \frac{1}{C''}\frac{\delta}{3}, \\
2C'\delta'\sqrt{\log \rho^{-1}} &= \frac{2C'}{C''}\frac{\delta}{3}, \\
\frac{CC'\delta'^2 \log\left(C_2' M \rho^{-1}\right) \log \rho^{-1}}{\log^5 N} &= \frac{CC'\delta^2 (\log(C_2' M) + \log \rho^{-1})}{9\left(C''\right)^2 \log^5 N} \leq \frac{CC'(\log(C_2' M) + \log \rho^{-1})}{3\left(C''\right)^2 \log^5 N}\frac{\delta}{3}.
\end{aligned}$$

As for the last term, if $\rho \geq \frac{1}{C_2' M}$, then $\frac{CC'(\log(C_2' M) + \log \rho^{-1})}{3(C'')^2 \log^5 N} \leq \frac{2CC' \log(C_2' M)}{3(C'')^2 \log^5 N} \leq \frac{2CC'}{3(C'')^2}$ (where we further supposed that $N \geq O(1)$). If $N^{-\log^4 N} \leq \rho \leq \frac{1}{C_2' M}$ (where the lower bound is from the theorem assumptions), then $\frac{CC'(\log(C_2' M) + \log \rho^{-1})}{3(C'')^2 \log^5 N} \leq \frac{2CC' \log \rho^{-1}}{3(C'')^2 \log^5 N} \leq \frac{2CC'}{3(C'')^2}$. By choosing C'' appropriately large, we then have

$$\mathbb{P}\left\{ \left|\!\left|\!\left| Z_1 \right|\!\right|\!\right| > \frac{\delta}{3} + \frac{\delta}{3} + \frac{\delta}{3} \right\} \leq 8\rho.$$

Putting the formula for δ' into $M \geq \frac{C_3 S \mu(\Psi)^2 \log^5 N}{\delta'^2}$ completes the proof. □

F.4 Change of RIP Conditioning with System Eigenvectors

In this section, we derive the worst case bound on how the RIP conditioning δ changes when the connectivity matrix W has non-orthonormal eigenvectors. The conditioning

of U can reduce the recoverable short term memory of the system by increasing the RIP conditioning value δ . If the Fourier matrix F has RIP conditioning

$$C\left(1-\delta\right)\left\|s\right\|_2^2 \leq \left\|Fs\right\|_2^2 \leq C\left(1+\delta\right)\left\|s\right\|_2^2$$

and U has a squared ratio of maximum and minimum singular values $\sigma_{\max}^2/\sigma_{\min}^2 = \gamma$, we can write the conditioning on the composite matrix UF in terms of the conditioning of F and the maximum and minimum singular values of U. We can find the upper bound via

$$\|UFs\|_2^2 \leq \sigma_{\max}^2\|Fs\|_2^2 \leq \sigma_{\max}^2 C(1+\delta)\|s\|_2^2,$$

and the lower bound via

$$\|UFs\|_2^2 \geq \sigma_{\min}^2\|Fs\|_2^2 \geq \sigma_{\min}^2 C(1-\delta)\|s\|_2^2.$$

We wish to consolidate these bounds into an RIP statement

$$C'\left(1-\delta'\right)\left\|s\right\|_2^2 \leq \left\|UFs\right\|_2^2 \leq C'\left(1+\delta'\right)\left\|s\right\|_2^2$$

where

$$\begin{aligned}
\sigma_{\min}^2 C(1-\delta) &= C'(1-\delta') \\
\sigma_{\max}^2 C(1+\delta) &= C'(1+\delta')
\end{aligned}$$

hold. These relationships can be used to solve for the RIP constants for the composite matrix:

$$\begin{aligned}
\delta' &= \frac{\frac{\gamma-1}{\gamma+1}+\delta}{1+\delta\frac{\gamma-1}{\gamma+1}} \\
C' &= \frac{1}{2}C\left(\sigma_{\max}^2+\sigma_{\min}^2+\delta(\sigma_{\max}^2-\sigma_{\min}^2)\right)
\end{aligned}$$

From these expressions, we can see that as the singular value spread of U diminishes (i.e. $\gamma \to 1$), the RIP conditioning does not change ($\delta' = \delta$). As γ grows, the conditioning factor grows accordingly.

REFERENCES

[1] Achlioptas, D., "Database-friendly Random Projections: Johnson-Lindenstrauss with Binary Coins," *J. Computer and System Science*, vol. 66, no. 4, pp. 671–687, 2003.

[2] Ailon, N. and Chazelle, B., "Approximate Nearest Neighbors and the Fast Johnson-Lindenstrauss Transform," in *Proc. 38th Annu. ACM Symp. Theory of Computing*, pp. 557–563, ACM, 2006.

[3] Ailon, N. and Liberty, E., "An Almost Optimal Unrestricted Fast Johnson-Lindenstrauss Transform," *Proc. ACM-SIAM Symp. Discrete Algorithms (SODA)*, pp. 185–191, 2011.

[4] Alon, N., "Problems and Results in Extremal Combinatorics , Part I," *Discrete Mathematics*, 2003.

[5] Anderson, G. D. and Qiu, S. L., "A Monotoneity Property of the Gamma Function," *Proc. American Mathematical Society*, vol. 125, no. 11, pp. 3355–3362, 1997.

[6] Andersson, J. and Strömberg, J.-O., "On the Theorem of Uniform Recovery of Structured Random Matrices," *arXiv preprint 1206.5986*, June 2012.

[7] Asefa, T., Kemblowski, M., Lall, U., and Urroz, G., "Support Vector Machines for Nonlinear State Space Reconstruction: Application to the Great Salt Lake Time Series," *Water Resources Research*, vol. 41, no. 12, 2005.

[8] Asif, S. M. and Charles, A. S., "Estimation and Dynamic Updating of Time-varying Signals with Sparse Variations," in *Proc. Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2011.

[9] Asif, S. M., Reddy, D., Boufounos, P. T., and Veeraraghavan, A., "Streaming Compressive Sensing for High-speed Periodic Videos," in *Proc. 17th IEEE Int. Conf. Image Processing (ICIP)*, pp. 3373–3376, IEEE, 2010.

[10] Ayaz, U. and Rauhut, H., "Nonuniform Sparse Recovery with Gaussian Matrices," *arXiv preprint 1007.2354*, July 2010.

[11] Bajwa, W. U., Haupt, J. D., Raz, G. M., and Nowak, R. D., "Compressed Channel Sensing," in *Proc. Conf. Information Sciences and Systems (CISS)*, (Princeton, NJ), Mar. 2008.

[12] Baraniuk, R. G., "More is Less: Signal Processing and the Data Deluge," *Science*, vol. 331, pp. 717–719, Feb. 2011.

[13] Baraniuk, R. G., Davenport, M. A., DeVore, R. A., and Wakin, M. B., "A Simple Proof of the Restricted Isometry Property for Random Matrices," *Constructive Approximation*, vol. 28, pp. 253–263, Jan. 2008.

[14] Baraniuk, R. G. and Steeghs, P., "Compressive Radar Imaging," in *Proc. IEEE Radar Conference*, pp. 128–133, 2007.

[15] Baraniuk, R. G. and Wakin, M. B., "Random Projections of Smooth Manifolds," *Foundations of Computational Mathematics*, vol. 9, no. 1, pp. 51–77, 2009.

[16] Baum, E. B., Moody, J., and Wilczek, F., "Internal Representations for Associative Memory," *Biological Cybernetics*, vol. 59, no. 4, pp. 217–228, 1988.

[17] Becker, S., Bobin, J., and Candès, E. J., "NESTA: A Fast and Accurate First-order Method for Sparse Recovery," *SIAM J. Imaging Sciences*, vol. 4, no. 1, pp. 1–39, 2009.

[18] Belkin, M. and Niyogi, P., "Laplacian Eigenmaps for Dimensionality Reduction and Data Representation," *Neural Computation*, vol. 15, no. 6, pp. 1373–1396, 2003.

[19] Berinde, R. and Indyk, P., "Sparse Recovery Using Sparse Random Matrices," *Proc. IEEE*, vol. 98, pp. 937–947, June 2010.

[20] Blumensath, T. and Davies, M. E., "Iterative Hard Thresholding for Compressed Sensing," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, 2009.

[21] Boufounos, P. T. and Asif, S. M., "Compressive Sampling for Streaming Signals with Sparse Frequency Content," in *Proc. Conf. Information Sciences and Systems (CISS)*, (Princeton, NJ), IEEE, 2010.

[22] Boufounos, P. T. and Asif, S. M., "Compressive Sensing for Streaming Signals using the Streaming Greedy Pursuit," in *Military Communications Conf. (MILCOM)*, (San Jose, CA), pp. 1205–1210, Ieee, Oct. 2010.

[23] Brin, M., Ott, W., and Yorke, J. A., "Enveloping Manifolds," *Topology and its Applications*, vol. 145, pp. 233–239, Nov. 2004.

[24] Brockwell, P. J. and Davis, R. A., *Introduction to Time Series and Forecasting*. Springer, 2002.

[25] Candès, E. J., "Compressive Sampling," in *Proc. Int. Congress of Mathematicians*, vol. 3, pp. 1433–1452, 2006.

[26] Candès, E. J., "The Restricted Isometry Property and its Implications for Compressed Sensing," *Comptès Rendus Mathématique*, vol. 346, no. 9-10, pp. 589–592, 2008.

[27] Candès, E. J. and Plan, Y., "A Probabilistic and RIPless Theory of Compressed Sensing," *IEEE Trans. Information Theory*, vol. 57, pp. 7235–7254, Nov. 2011.

[28] Candès, E. J. and Romberg, J. K., "Quantitative Robust Uncertainty Principles and Optimally Sparse Decompositions," *Foundations of Computational Mathematics*, vol. 6, no. 2, pp. 227–254, 2006.

[29] Candès, E. J. and Romberg, J. K., "Sparsity and Incoherence in Compressive Sampling," *Inverse Problems*, vol. 23, 2007.

[30] Candès, E. J., Strohmer, T., and Voroninski, V., "PhaseLift: Exact and Stable Signal Recovery from Magnitude Measurements via Convex Programming," *arXiv preprint 1109.4499*, 2011.

[31] Casdagli, M., Eubank, S., Farmer, J. D., and Gibson, J. F., "State Space Reconstruction in the Presence of Noise," *Physica D: Nonlinear Phenomena*, vol. 51, no. 1-3, pp. 52–98, 1991.

[32] Chandrasekaran, V., Recht, B., Parrilo, P. A., and Willsky, A. S., "The Convex Geometry of Linear Inverse Problems," *arXiv preprint 1012.0621*, 2010.

[33] Charles, A. S., Asif, S. M., Romberg, J. K., and Rozell, C. J., "Sparsity Penalties in Dynamical System Estimation," in *Proc. Conf. Information Sciences and Systems (CISS)*, (Baltimore, MD), IEEE, Mar. 2011.

[34] Charles, A. S. and Rozell, C. J., "Re-Weighted l_1 Dynamic Filtering for Time-Varying Sparse Signal Estimation," *arXiv preprint 1208.0325*, Aug. 2012.

[35] Charles, A. S., Yap, H. L., and Rozell, C. J., "Short-term Memory Capacity in Reccurent Networks via Compressed Sensing," in *Computational and Systems Neuroscience (Cosyne) Meeting*, (Salt Lake City, UT), 2012.

[36] Charles, A. S., Yap, H. L., and Rozell, C. J., "Short Term Network Memory Capacity via the Restricted Isometry Property," *Submitted*, May 2012.

[37] Chen, S. S., Donoho, D. L., and Saunders, M. A., "Atomic Decomposition by Basis Pursuit," *SIAM Review*, vol. 43, no. 1, pp. 129–159, 2001.

[38] Christensen, O., *An Introduction to Frames and Riesz Bases*. Birkhauser, 2003.

[39] Clarkson, K. L., "Tighter Bounds for Random Projections of Manifolds," in *Proc. 24th Annu. Symp. Computational Geometry*, pp. 39–48, ACM, 2008.

[40] Coifman, R. R. and Lafon, S., "Diffusion Maps," *Applied and Computational Harmonic Analysis*, vol. 21, pp. 5–30, July 2006.

[41] Coifman, R. R., Lafon, S., Lee, A. B., Maggioni, M., Nadler, B., Warner, F., and Zucker, S. W., "Geometric Diffusions as a Tool for Harmonic Analysis and Structure Definition of Data : Diffusion Maps," *Proc. National Academy of Sciences USA*, Jan. 2005.

[42] Dasgupta, S. and Gupta, A., "An Elementary Proof of the Johnson-Lindenstrauss Lemma," *Random Structures and Algorithms*, vol. 22, no. 1, pp. 60–65, 2002.

[43] Davenport, M. A., Boufounos, P. T., Wakin, M. B., and Baraniuk, R. G., "Signal Processing with Compressive Measurements," *IEEE J. Selected Topics in Signal Processing*, vol. 4, pp. 445–460, Apr. 2010.

[44] Davenport, M. A., Duarte, M. F., Wakin, M. B., Laska, J. N., Takhar, D., Kelly, K. F., and Baraniuk, R. G., "The Smashed Filter for Compressive Classification and Target Recognition," in *Proc. SPIE*, SPIE, Jan. 2007.

[45] Davenport, M. A., Hegde, C., Duarte, M. F., and Jan, L. G., "A Theoretical Analysis of Joint Manifolds," *Technical Report, Rice University*, 2009.

[46] Davenport, M. A. and Wakin, M. B., "Analysis of Orthogonal Matching Pursuit Using the Restricted Isometry Property," *IEEE Trans. Information Theory*, vol. 56, no. 9, pp. 4395–4401, 2010.

[47] De La Peña, V. H. and Giné, E., *Decoupling: From Dependence to Independence*. Springer Verlag, 1999.

[48] Dedieu, H., "Identifiability and Identification of Chaotic Systems Based on Adaptive Synchronization," *IEEE Trans. Circuits and Systems I*, vol. 44, no. 10, pp. 948–962, 1997.

[49] DeVore, R. A., "Deterministic Constructions of Compressed Sensing Matrices," *J. Complexity*, vol. 23, pp. 918–925, Aug. 2007.

[50] DeVore, R. A., Petrova, G., and Wojtaszczyk, P., "Instance-optimality in Probability with an l1-minimization Decoder," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 275–288, 2009.

[51] Diaconis, P. and Shahshahani, M., "On the Eigenvalues of Random Matrices," *J. Applied Probability*, vol. 31, no. 1994, pp. 49–62, 1994.

[52] Do Carmo, M. P., *Riemannian Geometry*. Birkhauser, 1992.

[53] Donoho, D. L., "Compressed Sensing," *IEEE Trans. Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[54] Donoho, D. L. and Grimes, C., "Image Manifolds which are Isometric to Euclidean Space," *J. Math. Imaging Computer Vision*, vol. 23, pp. 5–24, July 2005.

[55] Donoho, D. L. and Tanner, J., "Observed Universality of Phase Transitions in High-Dimensional Geometry, with Implications for Modern Data Analysis and Signal Processing," *Philosophical Transactions of the Royal Society A*, vol. 367, pp. 4273–4293, June 2009.

[56] Duarte, M. F., Davenport, M. A., Takhar, D., Laska, J. N., Sun, T., Kelly, K. F., and Baraniuk, R. G., "Single-pixel Imaging via Compressive Sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 83–91, 2008.

[57] Duarte, M. F., Davenport, M. A., Wakin, M. B., and Baraniuk, R. G., "Sparse Signal Detection from Incoherent Projections," in *Proc. Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, May 2006.

[58] Eftekhari, A., Yap, H. L., Rozell, C. J., and Wakin, M. B., "The Restricted Isometry Property for Random Block Diagonal Matrices," *arXiv preprint 1210.3395*, Oct. 2012.

[59] Eichhorn, R. and Linz, S. J., "Transformation Invariance of Lyapunov Exponents," *Chaos, Solitons and Fractals*, vol. 12, pp. 1377–1383, 2001.

[60] Figueiredo, M. A. T., Nowak, R. D., and Wright, S. J., "Gradient Projection for Sparse Reconstruction: Application to Compressed Sensing and Other Inverse Problems," *IEEE J. Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 586–597, 2007.

[61] Fraser, A. M. and Swinney, H. L., "Independent Coordinates for Strange Attractors from Mutual Information," *Physical Review A*, vol. 33, no. 2, pp. 1134–1140, 1986.

[62] Ganguli, S., Huh, D., and Sompolinsky, H., "Memory Traces in Dynamical Systems," *Proc. National Academy of Sciences USA*, vol. 105, pp. 18970–5, Dec. 2008.

[63] Ganguli, S. and Sompolinsky, H., "Short-term Memory in Neuronal Networks through Dynamical Compressed Sensing," in *Proc. Conf. Advances in Neural Information Processing Systems (NIPS)*, 2010.

[64] Gardner, W. A., *Introduction to Random Processes : with Applications to Signals and Systems*. New York: McGraw-Hill, 2nd ed., May 1990.

[65] Garnaev, A. Y. and Gluskin, E. D., "The Widths of an Euclidean Ball," in *Dokl. Akad. Nauk SSSR*, vol. 277, pp. 1048–1052, 1984.

[66] Ghomi, M. and Howard, R., "Tangent Cones And Regularity of Real Hypersurfaces," *arXiv preprint 1005.2761*, 2010.

[67] Gilbert, A. C. and Indyk, P., "Sparse Recovery Using Sparse Matrices," *Proc. IEEE*, vol. 98, no. 6, pp. 937–947, 2010.

[68] Grassberger, P. and Procaccia, I., "Measuring the Strangeness of Strange Attractors," *Physica D: Nonlinear Phenomena*, vol. 9, no. 1-2, pp. 189–208, 1983.

[69] Gray, A., "The Volume of a Small Geodesic Ball of a Riemannian Manifold," *The Michigan Mathematical Journal*, vol. 20, no. 4, pp. 329–344, 1974.

[70] Haimovich, A. M., Blum, R. S., and Cimini, L. J., "MIMO Radar with Widely Separated Antennas," *Signal Processing Magazine*, pp. 116–129, Jan. 2008.

[71] Haupt, J. D., Bajwa, W. U., Raz, G. M., and Nowak, R. D., "Toeplitz Compressed Sensing Matrices with Applications to Sparse Channel Estimation," *IEEE Trans. Information Theory*, vol. 56, no. 11, pp. 5862–5875, 2010.

[72] Haupt, J. D. and Nowak, R. D., "Compressive Sampling for Signal Detection," in *Proc. Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2007.

[73] Hegde, C., Wakin, M. B., and Baraniuk, R. G., "Random Projections for Manifold Learning: Proofs and Analysis," *Technical Report TREE 0710, Rice University*, 2007.

[74] Hinton, G. E., "Modelling the Manifolds of Images of Handwritten Digits," *IEEE Trans. Neural Networks*, vol. 1, no. 1, pp. 65–74, 1997.

[75] Horn, R. A. and Johnson, C. R., *Matrix analysis*. Cambridge University Press, 1990.

[76] Indyk, P. and Motwani, R., "Approximate Nearest Neighbors: Towards Removing the Curse of Dimensionality," in *Proc. 30th Annu. ACM Symp. Theory of Computing*, pp. 604–613, ACM, 1998.

[77] Isely, G., Hillar, C. J., and Sommer, F. T., "Deciphering Subsampled Data: Adaptive Compressive Sampling as a Principle of Brain Communication," in *Advances in Neural Information Processsing (NIPS)*, Oct. 2010.

[78] Jaeger, H., "The "Echo State" Approach to Analysing and Training Recurrent Neural Networks," *Technical Report GMD Report 148, German National Research Center for Information Technology*, 2001.

[79] Jaeger, H. and Haas, H., "Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication," *Science*, vol. 304, pp. 78–80, Apr. 2004.

[80] Johnson, W. B. and Lindenstrauss, J., "Extensions of Lipschitz Mappings into a Hilbert Space," in *Proc. Conf. Modern Analysis and Probability*, 1984.

[81] Kantz, H. and Schreiber, T., *Nonlinear Time Series Analysis*, vol. 7. Cambridge University Press, 2004.

[82] Kashin, B. S., "Diameters of Some Finite-dimensional Sets and Classes of Smooth Functions," *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya*, vol. 41, no. 2, pp. 334–351, 1977.

[83] Krahmer, F., Mendelson, S., and Rauhut, H., "Suprema of Chaos Processes and the Restricted Isometry Property," *arXiv preprint 1207.0235*, July 2012.

[84] Krahmer, F. and Ward, R., "New and Improved Johnson-Lindenstrauss Embeddings via the Restricted Isometry Property," *SIAM J. Mathematical Analysis*, vol. 43, pp. 1269–1281, Sept. 2011.

[85] Kugiumtzis, D., "State Space Reconstruction Parameters in the Analysis of Chaotic Time Series - the Role of the Time Window Length," *Physica D: Nonlinear Phenomena*, vol. 95, p. 20, Feb. 1996.

[86] LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D., "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, pp. 541–551, 1989.

[87] Ledoux, M., *The Concentration of Measure Phenomenon*, vol. 89. American Mathematical Society, 2001.

[88] Ledoux, M. and Talagrand, M., *Probability in Banach Spaces: Isoperimetry and Processes*, vol. 23. Springer, 1991.

[89] Li, J. and Stoica, P., "MIMO Radar with Colocated Antennas," *IEEE Signal Processing Magazine*, pp. 106–114, Sept. 2007.

[90] Lukoševičius, M. and Jaeger, H., "Reservoir Computing Approaches to Recurrent Neural Network Training," *Computer Science Review*, vol. 3, pp. 127–149, Aug. 2009.

[91] Lustig, M., Donoho, D. L., Santos, J. M., and Pauly, J. M., "Compressed Sensing MRI," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 72–82, 2008.

[92] Maass, W., Legenstein, R., and Bertschinger, N., "Methods for Estimating the Computational Power and Generalization Capability of Neural Microcircuits," in *Proc. Conf. Advances in Neural Information Processing Systems (NIPS)*, 2005.

[93] Maass, W., Natschläger, T., and Markram, H., "Real-time Computing Without Stable States: a New Framework for Neural Computation Based on Perturbations," *Neural Computation*, vol. 14, pp. 2531–60, Nov. 2002.

[94] Mallat, S., *A Wavelet Tour of Signal Processing: The Sparse Way*. Academic Press, 3 ed., 2008.

[95] Maus, A. and Sprott, J. C., "Neural Network Method for Determining Embedding Dimension of a Time Series," *Communications in Nonlinear Science and Numerical Simulation*, vol. 16, pp. 3294–3302, Aug. 2011.

[96] Mendelson, S., Pajor, A., and Tomczak-Jaegermann, N., "Uniform Uncertainty Principle for Bernoulli and Subgaussian Ensembles," *Constructive Approximation*, vol. 28, no. 3, pp. 277–289, 2008.

[97] Mongillo, G., Barak, O., and Tsodyks, M., "Synaptic Theory of Working Memory," *Science*, vol. 319, pp. 1543–6, Mar. 2008.

[98] Moon, T. K. and Stirling, W. C., *Mathematical Methods and Algorithms for Signal Processing*, vol. 13. Prentice hall New York, 2000.

[99] Muldoon, M. R., Broomhead, D. S., Huke, J. P., and Hegger, R., "Delay Embedding in the Presence of Dynamical Noise," *Dynamics and Stability of Systems*, vol. 13, no. 2, pp. 175–186, 1998.

[100] Nash, J., "C1 Isometric Imbeddings," *The Annals of Mathematics*, vol. 60, no. 3, pp. 383–396, 1954.

[101] Needell, D. and Tropp, J. A., "CoSaMP: Iterative Signal Recovery from Incomplete and Inaccurate Samples," *Communications of the ACM*, vol. 53, pp. 93–100, Mar. 2010.

[102] Niyogi, P., Smale, S., and Weinberger, S., "Finding the Homology of Submanifolds with High Confidence from Random Samples," *Technical Report No. TR-2004-08, University of Chicago*, pp. 1–23, Mar. 2006.

[103] Olshausen, B. A. and Field, D. J., "Emergence of Simple-cell Receptive Field Properties by Learning a Sparse Vode for Natural Images," *Nature*, vol. 381, no. 6583, pp. 607–609, 1996.

[104] Ott, E. and Yorke, J. A., "Controlling Chaos," *Physical Review Letters*, vol. 64, no. 11, pp. 1196–1199, 1990.

[105] Ott, W. and Yorke, J. A., "Learning About Reality From Observation," *SIAM J. Applied Dynamical Systems*, pp. 1–26, 2003.

[106] Park, J. Y., Yap, H. L., Rozell, C. J., and Wakin, M. B., "Concentration of Measure for Block Diagonal Measurement Matrices," in *Proc. Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, no. October, pp. 3614–3617, IEEE, Mar. 2010.

[107] Park, J. Y., Yap, H. L., Rozell, C. J., and Wakin, M. B., "Concentration of Measure for Block Diagonal Matrices with Applications to Compressive Signal Processing," *IEEE Trans. Signal Processing*, vol. 59, no. 12, pp. 5859–5875, 2011.

[108] Pecora, L. M. and Carroll, T. L., "Driving Systems with Chaotic Signals," *Physical Review A*, vol. 44, pp. 2374–2383, Aug. 1991.

[109] Qi, H. and Hughes, S., "Using the Kernel Trick in Compressed Sensing: Accurate Signal Recovery from Fewer Measurements," in *Proc. Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2011.

[110] Rauhut, H., "Circulant and Toeplitz Matrices in Compressed Sensing," in *Proc. Signal Processing with Adaptive Sparse Structured Representations (SPARS)*, vol. 9, Feb. 2009.

[111] Rauhut, H., "Compressive Sensing and Structured Random Matrices," in *Theoretical Foundation and Numerical Methods for Sparse Recovery*, 2010.

[112] Rauhut, H., Romberg, J. K., and Tropp, J. A., "Restricted Isometries for Partial Random Circulant Matrices," *Applied and Computational Harmonic Analysis*, vol. 32, pp. 242–254, Oct. 2012.

[113] Romberg, J. K., "A Uniform Uncertainty Principle for Gaussian Circulant Matrices," in *Proc. 16th Int. Conf. Digital Signal Processing (DSP)*, (Santorini), IEEE, July 2009.

[114] Romberg, J. K., "Compressive Sensing by Random Convolution," *SIAM J. Imaging Sciences*, vol. 2, no. 4, pp. 1098–1128, 2009.

[115] Roweis, S. T. and Saul, L. K., "Nonlinear Dimensionality Reduction by Locally Linear Embedding," *Science*, vol. 290, no. 5500, p. 2323, 2000.

[116] Rozell, C. J., Johnson, D. H., Baraniuk, R. G., and Olshausen, B. A., "Sparse Coding via Thresholding and Local Competition in Neural Circuits," *Neural Computation*, vol. 20, pp. 2526–2563, Oct. 2008.

[117] Rozell, C. J., Yap, H. L., Park, J. Y., and Wakin, M. B., "Concentration of Measure for Block Diagonal Matrices with Repeated Blocks," in *Proc. Conf. Information Sciences and Systems (CISS)*, (Princeton, NJ), IEEE, Mar. 2010.

[118] Rudelson, M. and Vershynin, R., "On Sparse Reconstruction from Fourier and Gaussian Measurements," *Communications in Pure and Applied Mathematics*, vol. 61, pp. 1025–1045, Aug. 2008.

[119] Sanandaji, B. M., Vincent, T. L., and Wakin, M. B., "Concentration of Measure Inequalities for Compressive Toeplitz Matrices with Applications to Detection and System Identification," in *Proc. IEEE Conf. Decision and Control (CDC)*, (Atlanta, GA), 2010.

[120] Sauer, T. D., Yorke, J. A., and Casdagli, M., "Embedology," *J. Statistical Physics*, vol. 65, pp. 579–616, Nov. 1991.

[121] Schreiber, T. and Kaplan, D. T., "Nonlinear Noise Reduction for Electrocardiograms," *Chaos (Woodbury, N.Y.)*, vol. 6, pp. 87–92, Mar. 1996.

[122] Shah, P. and Chandrasekaran, V., "Iterative Projections for Signal Identification on Manifolds: Global Recovery Guarantees," in *Proc. 48th Annu. Allerton Conf.*, (Allerton, CA), pp. 760–767, IEEE, Sept. 2011.

[123] Smola, A. J. and Schölkopf, B., "A Tutorial on Support Vector Regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199–222, 2004.

[124] Stark, J., "Delay Embeddings for Forced Systems. I. Deterministic Forcing," *J. Nonlinear Science*, vol. 9, pp. 255–332, June 1999.

[125] Strogatz, S. H., *Nonlinear Dynamics and Chaos: with Applications to Physics, Biology, Chemistry, and Engineering*. Westview Press, 1994.

[126] Takens, F., "Detecting Strange Attractors in Turbulence," in *Dynamical Systems and Turbulence, Warwick 1980* (Rand, D. and Young, L.-S., eds.), vol. 898 of *Lecture Notes in Mathematics*, pp. 366–381, Springer Berlin / Heidelberg, 1981.

[127] Talagrand, M., *The Generic Chaining*. Springer, 2005.

[128] Tang, G., Bhaskar, B. N., Shah, P., and Recht, B., "Compressed Sensing off the Grid," *arXiv preprint 1207.6053*, July 2012.

[129] Tenenbaum, J. B., de Silva, V., and Langford, J. C., "A Global Geometric Framework for Nonlinear Dimensionality Reduction," *Science*, vol. 290, pp. 2319–2323, Dec. 2000.

[130] Tillmann, A. M. and Pfetsch, M. E., "The Computational Complexity of the Restricted Isometry Property, the Nullspace Property, and Related Concepts in Compressed Sensing," *arXiv preprint 1205.2081*, May 2012.

[131] Tropp, J. A., "On the Linear Independence of Spikes and Sines," *J. Fourier Analysis and Applications*, vol. 14, no. 5, pp. 838–858, 2008.

[132] Tropp, J. A. and Gilbert, A. C., "Signal Recovery from Random Measurements via Orthogonal Matching Pursuit," *IEEE Trans. Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[133] Tropp, J. A., Laska, J. N., Duarte, M. F., Romberg, J. K., and Baraniuk, R. G., "Beyond Nyquist: Efficient Sampling of Sparse Bandlimited Signals," *IEEE Trans. Information Theory*, vol. 56, Jan. 2009.

[134] Turk, M. and Pentland, A., "Eigenfaces for Recognition," *J. Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–83, 1991.

[135] Uzal, L., Grinblat, G., and Verdes, P., "Optimal Reconstruction of Dynamical Systems: A Noise Amplification Approach," *Physical Review E*, vol. 84, July 2011.

[136] Vershynin, R., "Introduction to The Non-asymptotic Analysis of Random Matrices," in *Compressed Sensing, Theory and Applications* (Eldar, Y. and Kutyniok, G., eds.), ch. 5, pp. 210–268, Cambridge Univ. Pr., Nov. 2012.

[137] Wakin, M. B., *The Geometry of Low-dimensional Signal Models*. PhD thesis, Rice University, 2006.

[138] Wakin, M. B., Sanandaji, B. M., and Tyrone, V. L., "On the Observability of Linear Systems from Random, Compressive Measurements," in *Proc. IEEE Conf. Decision and Control (CDC)*, (Atlanta, GA), 2010.

[139] White, O. L., Lee, D. D., and Sompolinsky, H., "Short-term Memory in Orthogonal Neural Networks," *Physical Review Letters*, vol. 92, p. 148102, Feb. 2004.

[140] Whitney, H., "Differentiable Manifolds," *The Annals of Mathematics*, vol. 37, no. 3, pp. 645–680, 1936.

[141] Wright, S. J., Ma, Y., Mairal, J., Sapiro, G., Huang, T., and Yan, S., "Sparse Representation for Computer Vision and Pattern Recognition," *Proc. IEEE*, vol. 98, no. 1, pp. 1031–1044, 2010.

[142] Yap, H. L., Charles, A. S., and Rozell, C. J., "The Restricted Isometry Property for Echo State Networks with Applications to Sequence Memory Capacity," in *Proc. IEEE/SP Workshop Statistical Signal Processing (SSP)*, (Ann Arbor, MI), Aug. 2012.

[143] Yap, H. L., Eftekhari, A., Wakin, M. B., and Rozell, C. J., "The Restricted Isometry Property for Block Diagonal Matrices," in *Proc. Conf. Information Sciences and Systems (CISS)*, (Baltimore, MD), 2011.

[144] Yap, H. L. and Rozell, C. J., "Stable Takens' Embedding for Linear Dynamical Systems," in *Proc. IEEE Conf. Decision and Control (CDC)*, (Atlanta, GA), Dec. 2010.

[145] Yap, H. L. and Rozell, C. J., "On the Relation between Block Diagonal Matrices and Compressive Toeplitz Matrices," *Technical Report, Georgia Intstitute of Technology*, 2011.

[146] Yap, H. L. and Rozell, C. J., "Stable Takens' Embeddings for Linear Dynamical Systems," *IEEE Trans. Signal Processing*, vol. 59, pp. 4781–4794, Oct. 2011.

[147] Yap, H. L., Wakin, M. B., and Rozell, C. J., "Stable Manifold Embeddings with Operators Satisfying the Restricted Isometry Property," in *Proc. Conf. Information Sciences and Systems (CISS)*, (Baltimore, MD), 2011.

[148] Yap, H. L., Wakin, M. B., and Rozell, C. J., "Stable Manifold Embeddings with Structured Random Matrices," *arXiv preprint 1209.3312*, Sept. 2012.

[149] Yin, W., Osher, S., Goldfarb, D., and Darbon, J., "Bregman Iterative Algorithms for l1-minimization with Applications to Compressed Sensing," *SIAM J. Imaging Sciences*, vol. 1, no. 1, pp. 143–168, 2008.

[150] Zhu, M. and Rozell, C. J., "Population Characteristics and Interpretations of nCRF Effects Emerging from Sparse Coding," in *Computational and Systems Neuroscience (Cosyne) Meeting*, 2011.

[151] Ziniel, J. and Schniter, P., "Efficient High-Dimensional Inference in the Multiple Measurement Vector Problem," *Submitted*, Nov. 2011.